

# What if intelligence is inheritable?

**A gloomy view of modern society based on the assumption that those in charge form a "cognitive élite" which is likely to be dynastic mischievously begs too many questions.**

THE old chestnut of the social connotations of the hereditary components of intelligence (as measured by Intelligence Quotient, or IQ), has resurfaced again in the United States. The occasion is the impending publication of a book, *The Bell Curve: Intelligence and Class Structure in American Life* (Free Press, New York, 1994), by Charles Murray and the late Richard Herrnstein, a Harvard psychology professor. Last week the *New Republic*, regarded as a bastion of liberal opinion which nevertheless has a healthy distaste for political correctness, published an article summarizing the argument of the book together with a sheaf of often vitriolic expressions of dissent that will ensure that the months ahead will be filled with spuriously heated arguments.

When there are so many other things to argue about, that is a great misfortune. There are several things to say, not the least of which is that there should be no restraint on the publication of solidly based data bearing on the inheritance of IQ. The second is to remark on the error familiar from past rehearsals of this corrosive argument, that a demonstration of a genetic influence on IQ is in the broadest sense irrelevant to the determination of social policy in fields such as welfare, education and the relief of poverty. Murray and Herrnstein, somewhat gleefully, have blundered into that trap with their eyes open, driven by an apocalyptic view of what lies ahead for societies like that of the United States.

Whatever may be the deficiencies of IQ as a measure of intelligence, the index is an approximation of a kind. Accordingly, it is no surprise that individuals' IQ is correlated with their social success. Murray and Herrnstein make a great deal of their discovery that societies such as that of the United States are run by a "cognitive élite", but that is exactly what would be expected in a meritocracy. It is quite a different question to ask to what degree that meritocracy can be properly and prudently dynastic. There is ample anecdotal evidence, in the often gloomy records of hereditary dictatorships and family businesses, to suggest that the cognitive élite does not breed true. But that, of course, is what would be expected from a multigenic influence on intelligence, as represented by the Gaussian distribution (the 'bell curve') of measured IQ in sample populations.

Murray and Herrnstein do not, of course, suppose that intelligence or IQ is simply determined by people's genes. As they must, they allow for nurture as well as nature as important influences on the IQ of grown-up people. But they go on to argue that current social policies in the United States designed to remedy the disadvantages of the children of

those not belonging to the cognitive élite, loosely described as 'affirmative action', have failed in their intended consequences and should therefore be abandoned. The flaw is that the explanation may just as well be that the remedial measures so far taken are both insufficient and inappropriate.

Given the wealth of evidence in the past century of the malign effects of deprivation in the earliest months of childhood, affirmative action in recruiting young people to universities and colleges is at best a tardy palliative, which is not to say that it should be abandoned (as Murray and Herrnstein ask). Moreover, the provision of welfare food-stamps to the mothers of deprived children (a US practice) may be entirely insufficient to offset the other sources of deprivation in a child's emotional and intellectual environment. In short, the conclusion of the data Murray and Herrnstein have assembled could just as well be that there is an urgent need to find better ways of ridding young children of the cramping effects of deprivation, not simply for its own sake but for the intellectual strength of the society of which they will become a part.

Indeed, that is not a sufficiently prescriptive way of putting it. The document on which Murray and Herrnstein have laboured (by Murray's account, for eight years) will be welcomed by many of its readers for its illiberal conclusions, but the argument deserves to be turned on its head. If IQ is so convincingly correlated negatively with social disadvantage, and if at least some part of everybody's IQ is determined by nurture and not nature, should it not be a part of all governments' ambition to relieve the effects of social disadvantage and so make the best of their people's intellectual potential? The cognitive élite, at least, should recognize the force of that argument. □

## Europe after Finland

**The European Commission must exert itself in matching cultural diversity with economic unity.**

THE outcome of last Sunday's referendum in Finland on membership of the European Union (EU) will improve the chances that there will be 16 members next year. The vote (which has to be confirmed by a two-thirds majority of the Helsinki parliament) follows a similar result in Austria last June, but its immediate importance will be its effect on opinion in Sweden (to be tested by referendum on 13



November) and in Norway (where there is a vote two weeks later). If the Finnish vote (53:47 in favour of membership) had been more decisively in favour of membership, the effects elsewhere in Scandinavia would have been greater, but they are likely still to be positive. So the weekend's business in Finland is to be welcomed. What does that imply for the future of the EU, and for Europe as a whole?

That more means better is true in many respects. Because international trade is not a zero-sum game but a means by which Adam Smith's beneficent division of labour may be extended internationally, the EU as a single market will be strengthened to the benefit of all, or almost all. But supporters of the European enterprise insist that there is more to it than that. Nobody forgets that one purpose of the Treaty of Rome was to create a framework within which future European wars would be unthinkable. (The risk may have receded in the recent past, but there is still some way to go.) And then, as European statesmen are fond of saying, there is a cultural dimension that cannot be overlooked.

The trouble, sadly, is that the cultural argument conceals a difficulty. Europe is not culturally homogeneous, as can easily be told by the *gedanken* experiment of comparing a European journey from, say, Oslo to Naples with a journey of comparable distance from, say, Omaha (Nebraska) to Houston (Texas). By comparison with the United States, Europe is almost bewilderingly diverse in cultural respects. It is not merely that languages and school systems differ from one place to another, with the consequence that people's mobility is necessarily constrained, but that cultural diversity is part of Europe's interest and, ultimately, of its strength.

So how to preserve diversity while enlarging membership, perhaps ultimately as far as the Urals? The European Commission, often wrongly accused of homogenizing inclinations, should nevertheless be more active in this respect, if only to counter the impression created by the necessary uniformity of the directives it must repeatedly issue. But greater subsidies for itinerant folk-dancers will not suffice. The commission should instead concentrate on more efficient ways of teaching second and third languages to young people and on strengthening those parts of European culture where diversity can coexist happily with uniformity. Higher education and research are the obvious places at which to start (and the commission's existing schemes for making students and research fellows more mobile are a good beginning). But there is much more yet to be done.

The most urgent need is for a catalogue and then an understanding of Europe's system of higher education. Especially with the general drive for increased participation in higher education, national systems are steadily becoming more diverse. But they already differ from each other in quite radical ways; sometimes, major universities are managed centrally (Italy and France), sometimes regionally (Germany) and sometimes are supposedly autonomous, but centrally financed (Britain). Some employ academics from elsewhere on equal terms, others do not. Arrangements for the support of research differ from one country or even university to another. Links between universities and public research institutes are either strong or weak. To list these

differences is not to advocate their disappearance, although Europe's interest is that academics should be as free to wander the surface of the continent as they were in the fifteenth century. For the time being, it would help if the commission merely undertook to tell its constituents how diverse is the present system, and why. □

## Lady Luck's millennium

**Poor Britain's plans for the beginning of a new millennium will be almost literally in the lap of the gods.**

SINCE Britain decided last year on the creation of a national lottery, the government has been rubbing its hands with delight at the prospect of a cornucopia of off-budget income that can be spent on good causes for which the political case is only marginal — overseas aid charities, the domestic voluntary sector, the arts in general and the celebration of the millennium. On the face of things, the expectation is correct. The market-men have a naturally beguiling message: the puritanical can be tempted to buy lottery tickets on the grounds that they will be supporting good causes, the greedy by the promise of becoming millionaires.

The celebration of the millennium, financed by a fifth of the proceeds from the lottery (estimated to amount to £1.6 billion by 2000), is meant to take the form of the creation of "additional landmarks of very high quality". The government has appointed a Millennium Commission, with two members of the cabinet among its eight members, which in turn had appointed a chief executive, Mr Nicholas Hinton, to administer its funds. Hinton has a distinguished record as a charity administrator, most recently as director of the Save the Children Fund, going back over three decades. Then, last week, seven days before Hinton was due to start full-time in his new job, he was told that his services would not now be required. He has been compensated with three months' pay.

This is a curious way to carry on, to say the least of it. The nub of the disagreement is the commission's fear that Hinton would assume an unwelcome influence in the selection of projects qualifying for support. What this appears to mean is that Hinton would exclude the most wild and woolly proposals from formal consideration by the commission, and that he would also keep a zealous eye on the spending of the commission's funds. The commission's members appear to yearn for more direct involvement, which is their privilege.

The trouble is that the members are all part-timers, mostly with other demanding jobs. Worse, in the nature of the British government's way of dealing with ministers, its two placemen on the commission (Mr Stephen Dorrell and Mr Michael Heseltine) will have long since departed their jobs before the millennium is upon us, and so will not be called to account for how the money is spent. Worse still, the scale on which British people invest in lottery tickets has yet to be determined, but may be much smaller than the marketing consortium called Camelot expects. At this rate, the commission will soon be pleading for a postponement of the millennium. □



# Documents show French government is still looking to 'streamline' CNRS

**Paris.** Thousands of researchers from the Centre National de la Recherche Scientifique (CNRS), the largest basic-research organization in Europe, took to the streets across France last week to protest against a decision by Guy Aubert, the agency's new director-general, that researchers spend no more than 60 per cent of the funds they had been allocated for this year.

The protests were fuelled not only by the last-minute budget revision, but also by uncertainty over the future of the agency itself. One proposal contained in an unpublished government circular could reduce the number of CNRS scientific departments from

seven to three, for example, and its laboratories from 1,300 to as few as 400, with the control of many laboratories passing directly to universities.

From Marseilles to Orsay, last week saw a flurry of emergency meetings by researchers, amid an atmosphere of rebellion not seen since the last conservative government slashed research funding in 1986. Feelings were running particularly high as the government had already cut the research budget (excluding salaries) by 8 per cent in June (see *Nature* 369, 511; 1994).

Most researchers agree that CNRS needs to act to avert an imminent financial crisis.

But there has been outrage at the demand that they spend no more than 60 per cent of the funds they were allocated for 1994. Many have already spent this proportion, and therefore have no more money until the end of the year.

In a letter to Aubert, the five directors of the Institut Jacques Monod in Paris, which has already spent 75 per cent of its allocated budget for this year, demanded that he withdraw the revision. Its repercussions, they warn, would extend well beyond the end of the year.

Vice-chancellors and laboratory heads have similarly warned that the budget revision would "partially or totally paralyse" research. Scientists have said that this would mean that they will be unable to buy reagents to keep cell cultures alive or radioisotopes for DNA sequencing.

At Orsay, outside Paris, where several major CNRS laboratories stand to have FFfr20 million (US\$3.8 million) less than expected to spend — equivalent to running expenses for two months — researchers have said they will be forced to send postgraduate students to "work in the library" for the remainder of the year.

Aubert has said that it was "not easy" to take such measures so soon after being appointed, and that he understands that researchers may feel in a state of shock. "Previously oblivious to the financial situation, they have taken off their blindfolds, only to see the wall that is waiting for them," he said in an interview with *Le Monde*.

The "wall" in question is a FFfr550 million gap in the CNRS's finances which has accumulated over the past few years because annual payments from the state have failed to match promises made in the budget.

But trade unions argue that the CNRS has access to a 'reserve' of unspent funds that exceeds this deficit. They claim that Aubert is only taking such harsh action because he badly needs a "war-chest" to finance his proposed reforms of CNRS, and that CNRS has already taken out a bank loan to finance the proposed reforms.

Earlier this year, François Fillon, the minister for higher education and research, appeared to have abandoned wide-ranging reforms of the research system that the government had previously been contemplating, after strong resistance from researchers (see *Nature* 368, 675; 1994). But many of the main elements have now reappeared.

For example, an internal circular at the science ministry, dated 12 October, suggests reforms that would eventually transform the CNRS from a research organization directly employing thousands of ►

## 'Seaborgium' fails to win approval

**San Francisco.** Researchers at the Lawrence Berkeley Laboratory (LBL) in California will continue to call element 106 seaborgium, after Glenn T. Seaborg, the Nobel prizewinning chemist, even though the International Union of Pure and Applied Chemistry (IUPAC) has rejected the name. IUPAC has decided that it is not prepared to name new elements after living scientists.

The team that created and identified element 106 in 1974 announced its decision to name it after Seaborg, who shared in the discovery of plutonium and nine other transuranium elements, at a meeting of the American

Chemical Society (ACS) in March. The ACS's nomenclature committee subsequently accepted the name.

But seaborgium is not among the nine names recommended last month for elements 101 to 109 by IUPAC's Commission on Nomenclature of Inorganic Chemistry. The 20 chemists from 12 countries on the committee said they had decided elements should not be named after living people, and gave element 106 the name rutherfordium after Ernest Rutherford, the New Zealand physicist who discovered the atom.

"The majority of the commission felt it was necessary to have the perspective of history in relation to these discoveries before such a decision was made," the commission wrote in the announcement, to be published in the December issue of *Pure and Applied Chemistry*.

In general, the commission used names

proposed by the three major German, Russian and US groups involved in the nine discoveries — although not in the order suggested. But Albert Ghiorso, the senior nuclear scientist emeritus at Lawrence Berkeley Laboratory (LBL) who led the team that discovered element 106, said he felt that seaborgium had fallen victim to political horsetrading.

Ghiorso says that it was the first time that a discoverer of an element had not been allowed to have his or her name attached to the element concerned. He added that Seaborg was the obvious person to name the element after.

Seaborg himself, who won the 1951 Nobel prize in chemistry with E. M. McMillan for his work on the chemistry of the transuranium elements, says that he is disappointed by the decision, which he calls "illogical". He points out that some of his accomplishments took place more than 50 years ago, clearly withstanding IUPAC's test of time.

The IUPAC committee has proposed that elements 101, 102 and 103 be named mendelevium, nobelium and lawrencium respectively. It also recommended two Russian-proposed names of dubnium (after Dubna, near Moscow) and joliotium (after the physicists Frédéric Joliot) for elements 104 and 105.

All the proposed names are subject to ratification by the IUPAC Council at its meeting in Guildford, England, in August 1995.

**Sally Lehrman**

IMAGE  
UNAVAILABLE  
FOR COPYRIGHT  
REASONS

Jim Bechtle/Sandia National Laboratories

**Rutherford (left) to be honoured with element 106 in place of Seaborg (right).**



► scientists into a research council concerned primarily with distributing grants.

CNRS now gives research funds to its laboratories to spend as they see fit. But one proposal in the circular is that CNRS spend almost half its funds on a series of new research 'programmes', with funding awarded to individual research groups on the basis of peer review of grant proposals.

The circular claims that the goal of this is to improve the effectiveness and "dynamism" of research. But the trade unions argue that the main objective of the reforms is to shift control of research from the research organizations to the science ministry, and to reorient research towards political and economic goals.

Bernard Bigot, head of the science and technology commission at the research ministry, has said that one goal of the reforms is to strengthen the link between CNRS research and national objectives. Such objectives will be fixed by a 'Strategic Committee of Orientation' to be created within the science ministry next month.

CNRS has 27,000 staff, and a well-developed structure for implementing its research strategy. But, claim the unions, its political independence has become an obstacle to Fillon's plans to orient research towards national priorities. They describe the reforms as a "dilution of the CNRS's role, and an explosion of its structures".

The circular suggests that "administrative and scientific" control of more than half of CNRS laboratories — mainly those already 'associated' with universities — would pass to the universities, although theses would be free to apply for grants from the proposed CNRS programmes.

CNRS has yet to make any public comment on the proposed changes. But many scientists are concerned that universities, already hard-pressed to recruit the staff needed to keep pace with burgeoning student numbers, might not be able to afford to maintain new laboratories adequately.

The main concern, says Henri-Edouard Audier, a chemist at the Ecole Polytechnique, is that the government is being driven not primarily by considerations of research but by a desire to cut costs. By reducing the number of CNRS laboratories, it hopes to release funds for the planned strategic research programmes, he claims.

Furthermore, although France has many first-class research universities, not only in Paris but also, for example, in Lyons, Marseilles and Strasbourg, many are oriented primarily towards teaching, and lack the research traditions of their Anglo-Saxon counterparts. Some are also rife with nepotism, political wrangling and mediocrity.

Audier also claims that a proposal to reduce the number of CNRS scientific departments from seven to three — social sciences, life sciences and material sciences and engineering — is only a first step towards slimming CNRS down further.

**Declan Butler**

## Hewlett and Packard boost support for basic research

**San Francisco & Cambridge, UK.** The US computer and instrument company Hewlett-Packard last week became the source of two separate grants aimed at supporting basic research, acknowledging its founders' belief in the past (and future) contribution of such research to their company's fortunes.

In California, William R. Hewlett and David Packard, two of the original Silicon Valley pioneers, announced a joint gift of \$77.4 million to Stanford University, where they first met in the 1930s. The gift — the single largest in Stanford's history — completes the financing needed for a \$175-million, nine-building science and engineering complex.

In a separate development, the president of Britain's Royal Society, Sir Michael Atiyah, formally inaugurated a new Basic Research Institute in the Mathematical Sciences (BRIMS), based at Hewlett-Packard's research laboratories in Bristol.

BRIMS will have close links with the Isaac Newton Institute in Cambridge, of which Sir Michael is the director. The company will also finance a fellowship at the Newton Institute, funding its first full-time staff member.

The gift to Stanford will pay for the construction of four major buildings, as well as the demolition of several existing ones to make room for them. The project began in the early 1980s with a pledge of \$40 million by Hewlett and Packard. The complete design is centred on a wooded courtyard and is intended to foster collaboration between disciplines, according to David Glen, principal gift director at Stanford.

"We've tried to look at the way our scientists do science and to design a campus that meets their needs — not only for science, but also for teaching," says Glen. He adds that the proximity of the scientific laboratories to humanities and business would help to realize Hewlett and Packard's vision of graduates who have a broad understanding of a variety of disciplines.

An electrical engineering building, together with the new Gates Information Science Building — made possible by a \$6-million donation from Bill Gates of Microsoft Corporation — and the expansion of the Center for Integrated Systems, will combine research and studies in both computer hardware and software.



**Packard: eyes on long-term strategy.**

A laboratory for advanced materials research will allow chemists to work closely with engineers, while a statistics department building will be placed next to a new complex of lecture halls and classrooms. The entire new section will be close to the university's medical centre and biology buildings.

In a joint statement, Hewlett and Packard, who were both members of Stanford's class of 1934, said: "We believe this gift will ensure that Stanford University will have leadership in science and engineering second to none during the 21st century."

The new UK institute is intended to give the company's researchers in Europe and the United States access to advanced mathematical thinking — particularly in areas such as nonlinear mathematics — which is likely to provide the underpinning of future developments in fields ranging from quantum devices to broadband networking.

According to John Taylor, the director of the HP Laboratories in Bristol at which BRIMS will be based, the institute is being financed from money that the company has set aside at the personal direction of David Packard to support basic research. Packard stepped down as chairman in September but remains chairman emeritus.

"Last year, Packard said that it would be okay to spend some resources on long-range topics, and that has given us permission to resist pressure from the rest of the company and work on things that have no particular application at the present," says Taylor.

At the same time, Taylor points out that non-linear mathematics has potential relevance to many parts of the company's business, such as soliton propagation in optical fibres, the chaotic modelling of cardiac arrhythmia and the statistical mechanics of very large networks.

The Newton Institute, which was opened two years ago, will provide support for short courses on advanced mathematics at Bristol and will also reinforce the activities of BRIMS in other ways. In return, the company has provided the institute with 10 new computers, as well as financing the research fellow in Cambridge and another shared appointment with the University of Bristol.

Atiyah says that Hewlett-Packard's backing for the Cambridge institute will have no direct influence on its choice of research topics. At the same time, he is keen to emphasize that, with university researchers being urged to contribute towards wealth creation, the new deal illustrates how even basic mathematics can play its part. "It is in the spirit of last year's white paper on science and technology," he says.

**Sally Lehman & David Dickson**

# Research must precede clean-up, DoE warns

**Boston.** A leading official of the US Department of Energy (DoE) warned last week that the department needs to impose a major change in the direction of its \$3-billion-a-year research programme if it is to find a way of clearing up the legacy left by its nuclear weapons research.

But he added that the department's laboratories, which are currently searching for a new mission, are not necessarily the right place for conducting new research into clean-up techniques, as they are too expensive.

There is also evidence that the public is growing impatient with the government's approach to clean-up. This has been reflected in a recent congressional cut of \$400 million in the Department of Defense's \$2.6 billion clean-up budget for the 1995 financial year, which started this month.

Tom Grumbly, assistant secretary in charge of the clean-up, was addressing a meeting at the Massachusetts Institute of Technology (MIT) on efforts being made by the Department of Defense and Department of Energy. He said that he intends to resist pressure to increase the speed with which contaminated sites are being tackled, preferring to wait until the research has been done to find the right means of doing so.

"We want to titrate the amount of clean-up that we do with the amount of new technology we have coming through," Grumbly said. Clean-up of department sites, which include the entire production chain for US nuclear weapons, would take "not 30 years, but a lifetime," he added. It would require much fundamental research on ways of dealing with chemical and radiation problems.

Calling on scientists to join in the search for new clean-up techniques, Grumbly referred back to the early nuclear work of Niels Bohr, Robert Oppenheimer and Edward Teller. "Their legacy is not finished," he said. Today's scientists have an obligation "to close the circle and apply the same scientific vigour" to the clean-up problem, he said.

Until five years ago, the DoE spent no money on research into clean-up techniques. Such work now accounts for about \$400 million of the department's \$3 billion annual research and development budget.

Further expansion will have to come at the expense of other research areas, such as alternative energy sources (also an administration priority), high-energy physics, or nuclear weapons. The last option is more problematic that it sounds, as the US nuclear weapons strategy is to cease production and tests, but maintain capability on the basis of superior research and development.

The department says it is making substantial progress on the clean up of the 160 contaminated sites left by the nuclear weapons programme. "In one year, the programme has gone from being a total failure to at least

offering hope," says Grumbly. Of \$2 billion budgeted for environmental restoration, half goes on "actual remediation" and the rest, he concedes, on the "analysis-paralysis part of the programme", under which the department assesses problems, haggles with regulators and then reassesses them.

From 1 January, Grumbly will inherit



full responsibility for nearly all of these sites and their mission will switch from being available to produce weapons to cleaning themselves up. The budget of his office will account for about \$7 billion out of the department's \$12 billion total, and clean-up will displace the nuclear weapons programme as the DoE's largest responsibility.

The Department of Defense has an equally large non-nuclear clean up plan, part of which has been speeded up under its base closure programme, which allows small, site-specific teams to manage 'fast-track' clean-up. The teams — which each include a military coordinator, an official from the Environmental Protection Agency and a state regulator — have special powers to ensure that plans are carried out.

But the DoE has no such short cut in prospect. All it can do is hope that Congress will have the patience to pay for the research that will eventually allow the nuclear mess to be cleaned up over many decades.

At the same time, several speakers at the MIT conference voiced scepticism about the department's central strategy of winning back public trust, as it is trying to do through an 'openness' campaign led by Hazel O'Leary, the energy secretary.

Todd LaPorte, a professor of politics at the University of California at Berkeley, who is preparing a report for O'Leary on the question of nuclear waste and public trust, reports that "the public has no trust in the department".

Even Grumbly hinted that neither he nor O'Leary is immune to the idea that the Energy Department, as it currently operates, may not be up to the task in hand. "I have serious doubts, some days," that the department can do the clean-up, he says. "And I know the secretary does too."

Colin Macilwain

## Cable set to confound HTS critics

**Boston.** High-temperature superconductors move closer to real-world application to power transmission this week with the launch of a \$6-million project to build a prototype underground power line.

The project is being mounted jointly by the Electric Power Research Institute (EPRI) — a power industry research collaboration — American Superconductor Corporation of Westborough, Massachusetts, cable company Pirelli, and the laboratories of the US Department of Energy.

Details of the project were announced in Boston by energy department officials and Senator Edward Kennedy (Democrat, Massachusetts). If it fulfils its objective, the project will lead to the economically feasible replacement of existing underground copper power lines in urban areas with lines based on high-temperature superconductivity (HTS).

Success would also confound critics who said the ceramic materials, invented in 1987 by IBM researchers in Switzerland, were too fickle in their conductivity, and too brittle, for such use.

EPRI, which will fund almost half of

the project, sees the 30-metre prototype as a step to upgrading and replacing 2,200 miles of ageing copper trunk power lines under the streets of US cities. By replacing these *in situ* with superconducting lines of two to five times the power-handling capacity, power companies could save the costs of re-excavating the lines.

"They've borrowed the retrofit concept from the telecommunications industry, which has been pulling out copper links and putting in optical fiber," said Greg Yurek, a former materials professor at the Massachusetts Institute of Technology who now heads American Superconductor.

The prototype power line will consist of a matrix of superconductors wound around a stainless steel tube, through which liquid nitrogen will flow to keep them cool. Paul Grant of EPRI says this is not seen as a disadvantage, as existing underground copper cables have to be cooled by oil. The challenge in building the prototype will be to produce and machine-wind the superconductor at sufficient quality and satisfactory cost. C. M.



# Neutron scattering reaps Nobel for physics

**London.** Techniques for probing the structure and properties of matter on the atomic scale have found much favour among the Nobel committee, from the Braggs' award for X-ray diffraction in 1915 to that for the scanning tunnelling microscope in 1986. This year's physics prize extends the same theme through its joint award to Bertrand Brockhouse and Clifford Shull for their work in developing neutron inelastic scattering and neutron diffraction respectively.

Both techniques rely on the scattering of neutrons by matter. Neutron scattering has never been the cheapest of experimental techniques. But the pervasive role of neutron scattering in condensed-matter physics can be justified on the grounds that it provides information that is otherwise inaccessible.

Neutron diffraction works in the same way as X-ray diffraction; the fact that the quantum-mechanical (de Broglie) wave-

IMAGE  
UNAVAILABLE  
FOR COPYRIGHT  
REASONS

## Shull: preparing to celebrate.

length of thermal neutrons is of the same order as the wavelength of X-rays means they probe long-range order on the same length scales, which are those typical of atomic lattices. But neutron diffraction differs in important respects stemming from how neutrons interact with matter.

The neutron-scattering power of atoms depends in a very erratic way on the atomic

mass. Neutron diffraction can therefore distinguish elements that are invisible or indistinguishable to X-rays. Hydrogen in particular has a large neutron scattering cross-section, and neutrons can be used to find these light atoms in a crystal structure.

Adjacent elements in the periodic table are hard to differentiate with X-rays, but can have vastly different neutron scattering cross-sections. Neutron diffraction has therefore revealed the structure of metal alloys such as brass (a copper-zinc alloy). And even isotopes of the same element can look very different to neutrons. Furthermore, as neutrons have a magnetic moment, they can be scattered by magnetic atoms.

Working at Oak Ridge in Tennessee with the group of E. O. Wollan, Shull played a central role in the development of instrumentation for routine crystallographic analysis with neutrons. He was the first to demonstrate magnetic diffraction, from the compound manganese oxide. Before his retirement Shull was working at the Massachusetts Institute of Technology (MIT).

Collisions between neutrons and atoms can also be inelastic, leading to transfer of energy. Neutron scattering can therefore be used as a spectroscopic technique. In the 1950s, Brockhouse developed the pioneering neutron spectrometer that forms the basis of instruments still in use today — an achievement that he carried off almost singlehandedly, according to Bob Birgeneau of the MIT.

Inelastic neutron scattering is now a mainstay of physicists studying the dynamics of condensed phases, such as the lattice vibrations of solids (known as phonons) and the molecular motions in adsorbed phases. Brockhouse was the first to measure the phonon dispersion curve of a solid, a spectrum of the lattice vibrational energies. He has now retired from the physics department of McMaster University in Canada.

The selection of Brockhouse and Shull as the founders of neutron scattering is "absolutely the right choice", says Jim Jorgensen of Argonne National Laboratories. "They did amazing things in the early days," he adds. Peter Day of the Royal Institution in the United Kingdom and ex-director of the Institut Laue-Langevin in Grenoble, France, Europe's premier neutron-scattering centre, says that "these are giants of the field".

Although Europeans have been well-served for neutron studies, the US situation remains under a cloud. In particular, the future of the proposed \$2.7-billion advanced neutron source at Oak Ridge is uncertain. "We're in a sad state of disarray in neutron funding in the United States," says Jorgensen. Will the Nobel awards boost the status of neutron scattering in condensed-matter science? "I'm counting on it," says Birgeneau.

**Philip Ball**

## Chemist honoured for 'carbocations'

**London.** The products of organic chemical reactions, such as ferrocene, dienes and crown ethers, have provided many Nobel prizes for chemistry in the past. This year, in contrast, the prize has been awarded for work on the transient species that come and go along the reaction pathway.

The Swedish Academy last week awarded the prize to Hungarian-born chemist George Olah for his studies of carbocations, positively charged species that are a common intermediate of many organic reactions.

Olah joined the scientific exodus from Hungary following the uprising of 1956. He joined the Dow Chemicals Company in Canada in 1957 and is now director of the Loker Hydrocarbon Research Institute at the University of Southern California in Los Angeles.

There are two ways of turning a carbon-based molecule into a carbocation. One is to remove an anion, leaving a carbenium ion with a formally three-coordinate carbon atom. The other is to add a hydrogen ion to a saturated carbon, creating a carbonium ion with a five-coordinate carbon.

In the 1920s and 1930s, C. K. Ingold proposed that carbenium ions are ephemeral reaction intermediates of certain elimination and substitution reactions of organic compounds, in which the first step is the loss of an anion. So common are these reactions that carbenium ions "dominate half of organic chemistry in terms of how things happen", according to Philip Eaton of the

University of Chicago.

Carbonium ions, formed by protonation of alkyl carbons using superacids such as  $\text{HSO}_3\text{F}$  and  $\text{HF-SbF}_5$ , are of particular interest because their bonding configuration cannot be rationalized according to classical ideas of valence, under which the positive charge is attached to an atom with a well-defined coordination number. Rather, these ions are 'non-classical' structures in which bonding is delocalized between several atoms, with the positive charge distributed among them.

In the days before ultrafast spectroscopic techniques, the lifetimes of these unstable species were too short for them to be observable. Working at Dow in the 1960s, Olah found a way to stabilize carbenium ions for long enough to use nuclear magnetic resonance and electron spectroscopy (ESCA) to elucidate their structure and properties.

Olah used the Lewis superacid  $\text{SbF}_5$  to extract halide ions from alkyl halides in solvents that were such poor nucleophiles that they did not readily attack the positively charged species. When he announced his success in analysing long-lived carbenium ions at a conference at Long Island in 1962, the news "hit like a bombshell" says Paul von Ragué Schleyer of the University of Erlangen-Nürnberg, Germany, who was there to hear it.

Schleyer points to the practical value of Olah's work as a major justification for the award. Protonation of hydrocarbons may lead to isomerization, for example converting straight-chain hydrocarbons into branched ones, which boosts the octane rating of hydrocarbon fuels. Such conversions are now commonly carried out using solid superacids such as zeolites. **Philip Ball**



**George Olah**



# Europe targets Mercury for future mission

**Paris.** The European Space Agency (ESA) announced this week that its priorities in space science early next century will probably include a mission to the planet Mercury and an interferometric study of the distances, motions and luminosities of tens of millions of stars in the Milky Way.

The two missions were chosen from 110 proposals submitted by European scientists — and 14 from scientists in the United States — as candidates for 'cornerstone' missions in ESA's planned Horizon 2000-plus space science programme. This will run for ten years from 2006, and will follow the current Horizon 2000 programme.

The quality and cost-effectiveness of ESA's 'Horizon' programmes have been widely acclaimed. Many attribute their success to the funding stability they enjoy; ESA's 13 member states are obliged to contribute a fixed proportion of their gross national product, while future funding levels are agreed in five-year chunks. This year, ESA will spend ECU381 million (US\$472 million) on science, about one-tenth of its total budget.

'Cornerstone' missions are each budgeted to cost around ECU625 million. The Mercury mission would probably include an orbiting observatory and also a landing craft. According to Roger Bonnet, head of science at ESA, the US National Aeronautics and Space Administration (NASA) has offered to take part if ESA decides to go ahead. As a result, the mission could become a European-led international effort.

The astronomy cornerstone would use interferometry to survey stars in the Milky Way, to search for Jupiter-sized planets — and signs of life — around stars in our Galaxy, and to test the theory of general relativity. Instruments on the craft would increase the resolution of current satellites by one to two orders-of-magnitude, giving them a capacity to resolve from Earth objects two millimetres apart on the Moon.

In what would be ESA's first venture into fundamental physics, the space agency is also considering a cornerstone mission to detect gravitational waves emitted, for example, from black holes, binary stars and

the early Universe. Such a mission would also be technologically daunting, requiring the launch of six satellites millions of kilometres apart, and keeping them connected by laser beams.

But much depends on funding. Bonnet points out that, although the first two missions cornerstones could be achieved within a static budget, one to detect gravitational waves would require a 4 to 5 per cent increase in ESA's science budget between 2000 and 2005. ESA will carry out technological studies before deciding whether to request further funding.

In addition to the big cornerstone missions, Horizon programmes also include several 'medium-sized' missions, budgeted at around ECU345 million, chosen by in open competition combined with peer review. One candidate for Horizon 2000-plus is a mission to Mars. ESA would have proposed this as a cornerstone, says Bonnet, if the United States, Japan and Russia did not already have similar plans.

ESA officials say that solar missions should also be included as medium-sized missions in Horizon 2000-plus, although specific details have not yet been decided. Looking beyond Horizon 2000 plus, the survey committee has recommended that Europe build on its lead in X-ray, gamma-ray and infrared astronomy.

ESA's science programme committee is

IMAGE  
UNAVAILABLE  
FOR COPYRIGHT  
REASONS

**The Red Planet: ESA's new priority.**

expected to approve the proposals next month. Final approval will have to wait until the meeting of Europe's space ministers late next year, which is expected to look closely at the balance between science missions and the more expensive (and increasingly criticized) efforts in manned spaceflight. Ministers will not need to commit money to Horizon 2000-plus until 2000, says Bonnet, as preliminary work can be done within existing budget allocations. **Declan Butler**

## Germany holds up bioethics treaty

**Paris.** The parliamentary committee of the 32-member Council of Europe last week postponed a vote on the draft European Convention on Bioethics after opposition from Germany. German representatives on the committee argued that the text should be more restrictive and less ambiguous.

In particular, the German government wants the council to support tighter regulations on the use of human embryos in research. Its opposition has highlighted the difficulty in achieving international consensus on bioethics issues; the provisions of a draft released in July already amounted to

little more than the lowest common denominator (see *Nature* 370, 3; 1994).

But Germany feels that a provision of the draft convention forbidding research on human embryos more than 14 days old is insufficient. It also said that provisions on handicapped people could open the way to their use as research subjects.

Supporters of the convention claim that it provides an umbrella of fundamental principles, and that member states can implement stricter legislation if they choose.

German resistance stems partly from the fact that the bioethics convention became a political issue in the run-up to the general elections last Sunday (16 October), largely as a result of a general sensitivity in Germany to issues related to genetics and human experiments. All the political parties had promised they would reject the text on the grounds that it did not go far enough.

The committee had hoped to approve the text ahead of the next meeting in January 1995. It will now seek a compromise with Germany to vote in at the January meeting. If this goes to plan, the convention — which is now two years behind schedule — may be presented to member states for ratification in March next year. **D. B.**

## Berlin protestors condemned by minister

Munich. Scientists at east Berlin's large teaching hospital, the Charité, upset at a planned merger with west Berlin's Rudolf Virchow clinic, have been sharply criticized by Berlin's ministry of research for bringing the issue into the recent general election campaign.

A group of scientists representing non-professional scientific staff leafleted the city last week, calling on voters not to vote in last Sunday's election for the two major parties in Berlin's grand coalition,

the Social Democrat Party (SPD) and the Christian Democrat Union (CDU), because of the damage they claim that they have caused the hospital.

Manfred Erhardt, Berlin's research minister, condemned the attempt to use the hospital's problems in an election campaign as "irresponsible and damaging to the Charité". But supporters of the scientists claim they had been left with no option, as no officials appeared to be listening to their complaints. **Alison Abbott**

# 'Steel' university tops the ratings in South Korea

**Seoul.** South Korea's new Pohang University of Science and Technology (POSTECH) has come top in the country's first attempt to rank its universities. POSTECH is supported by one of the world's most profitable steel companies, Pohang Iron and Steel Company (POSCO).

The ranking was made on the basis of various factors related to teaching and research. On most of these POSTECH came out ahead of the Seoul National University (SNU), which has traditionally been considered South Korea's best university. The survey of 131 universities was carried out at the end of last month by the *Joong-ang Daily News*, a leading national newspaper. It was followed a week later by a similar government report to the National Assembly.

But both analyses have one glaring omission, namely the exclusion of the Korea Advanced Institute of Science and Technology (KAIST), another top-ranking institution. Unlike the other universities, which come under the Ministry of Education, KAIST belongs to the Ministry of Science and Technology (MOST). As such, it is considered an outsider by other universities and the education ministry.

POSTECH came out top because of a low student/faculty ratio (6.4 to one, about a third that of SNU), a large budget for research and education (a significant portion of which comes from POSCO), a well-equipped library and laboratories and a large amount of floor space.

SNU has more publications cited in the Science Citation Index (SCI) of the US Institute for Scientific Information than POSTECH. But POSTECH's president, Sooyoung Chang, points out that the average number of SCI publications per faculty member in his institution is higher than that at SNU.

SNU leads the newspaper survey in terms of the number of graduates who have passed national examinations for professions such as the civil service, law and accounting, or who are chief executives of companies. POSTECH does not appear near the top in these categories. But it has been in existence for only eight years and does not cover all the subject areas of the national examinations.

KAIST, which is located in Taeduck science town in central South Korea, would rank at or near the top in many of the categories in the survey. Since the beginning of 1992, for example, its researchers have produced almost a quarter of all South Korean publications listed in the SCI.

It also has a similar annual budget to POSTECH (US\$60 million). The US Accreditation Board for Engineering and Technology reported in an external review of the university in 1992 that KAIST had the potential to become "one of the top institutions in the world" (see *Nature* 364, 379; 1993).

KunMo Chung, president of the Institute for Advanced Engineering and a former minister of science and technology, says that KAIST was probably omitted from the survey to avoid embarrassing SNU. There has been rivalry between the two universities ever since KAIST was established by MOST in 1971 as South Korea's first graduate university. The decision was made over the objections of both SNU and the education ministry, which had wanted to strengthen graduate education at SNU.

Sang Chul Shim, the president of KAIST, says he is not surprised by the absence of his university from the ranking exercise — he is often not invited to official meetings of South Korea's university presidents.

But whatever the reason for KAIST's exclusion, it underlines a sharp division between MOST and the Ministry of Education over the administration of research. KAIST receives generous financial support from the government through MOST, while most South Korean universities are starved of government funds.

POSTECH is an exception because of the generous support it receives from POSCO. The company provides about half of the university's research and development funds in the form of grants for project research and also supplied most of the construction costs of the university (see *Nature* 364, 380; 1993).

David Swinbanks

# HIV vaccines get the green light for Third World trials

**Geneva.** An advisory committee to the World Health Organization (WHO) has recommended that large-scale Phase III trials of HIV vaccines should be allowed to proceed in developing countries. The decision was made even though the United States has not given permission for the trials of two such candidate vaccines, made from genetically engineered forms of HIV's envelope protein, primarily because of doubts over their effectiveness.

The recommendation was made at a meeting in Geneva last week of 40 scientists, public-health experts and others concerned with the growing spread of AIDS in developing countries. Under normal circumstances, WHO would be reluctant to support trials elsewhere of a vaccine that had not passed safety trials in its country of origin.

In the current situation, however, the advisory group eventually agreed that the different conditions prevailing in many developing countries justified its recommendation to proceed with the trials. But it also emphasized that "there must be no ethical or scientific short cuts to speed the process", and that the trials must be approved by the appropriate national authorities.

Even though the decision was unanimous, several individual committee members are believed to have reservations about the outcome. Their concern was focused on the possibility that the trials could unfairly raise expectations of success, and in some cases fears that vaccines could actually enhance disease in those who might later become infected. But others argued that even a vaccine with a relatively low level of efficacy might save large numbers of lives.

A key candidate country for future experiments is Thailand, where the US Department of Defense is thought to be planning trials of a recombinant gp120 vaccine — one of the two vaccines under consideration for trials in the United States — in collaboration with the US company Biocine. Genetech is also planning trials of its own recombinant gp120 vaccine in a Third World country, but no definite plans have yet been announced by either company.

In a joint statement issued at the end of the meeting, the participants said that, although the scientific considerations for commencement of HIV efficacy trials should be the same in all parts of the world, "it is recognized that public health needs, as well as feasibility considerations, may be different from one country to another".

As a result, the group says, an analysis of the various types of considerations that should be made in relation to trials with specific candidate vaccines "may result in different recommendations in relation to efficacy trials in different parts of the world". □

## Ranking of South Korea's universities

|                                | RANK    |          |                  |
|--------------------------------|---------|----------|------------------|
|                                | 1st     | 2nd      | 3rd              |
| Student/faculty ratio          | POSTECH | Kong Ju  | SNU              |
| Research funds/professor       | POSTECH | SNU      | Nat. Fish. Pusan |
| Education outlays/student      | POSTECH | Hallym   | Dan Kook         |
| Library purchases/student      | POSTECH | Hallym   | Sae Jong         |
| Floor space                    | POSTECH | Sung Sim | Chung Ang        |
| Research facilities            | POSTECH | Myong Ji | Hallym           |
| SCI publications               | SNU     | POSTECH  | Yonsei           |
| Passes in national board exams | SNU     | Korea    | Yonsei           |
| Chief executives in companies  | SNU     | Korea    | Yonsei           |

POSTECH: Pohang University of Science and Technology.

SNU: Seoul National University.

Nat. Fish. Pusan: National Fisheries University of Pusan

The Korea Advanced Institute of Science and Technology (KAIST), was excluded from this analysis

Source: *The Joong-ang Daily News*



# Protein gene patent faces challenge in court

**Munich.** Amgen's European patent of the gene responsible for the protein erythropoietin is being challenged by Genetics Institute (GI) on the grounds that it lacks the necessary 'inventive step'. GI says that Amgen was able to discover the gene because it had access to the only supplies of the human form of the protein.

The claim is one of several being made in the final chapters of legal wranglings between the Californian-based Amgen and GI, based in Boston, over the rights to erythropoietin. The protein, which stimulates the production of red blood cells from bone marrow and is used to treat anaemias, has an estimated world market of US\$2 billion and a European market of US\$500 million.

Last month, the European Patent Office appeals board heard evidence on a series of challenges launched against Amgen by GI and five other companies. The board took the unusual step of delaying the ruling on the grounds that the issues involved were too complex for an immediate decision. Its verdict is expected on 21 November.

The dispute between GI and Amgen over erythropoietin has been long and bitter. GI was first granted a US patent for a pharmaceutical preparation of the protein in a high-activity homogeneous form in 1987. A European patent followed in 1992.

GI was also engaged in a close-run race with Amgen to clone the erythropoietin gene. The task was complicated by the difficulties of obtaining a pure source of the protein for cloning. Erythropoietin is produced in sig-

nificant quantities only in patients suffering from anaemia, when the body tries to rebuild its damaged red blood cell count.

Such patients are usually treated immediately. But Amgen secured supplies of erythropoietin extracted from the urine of three untreated Japanese patients suffering from aplastic anaemia, an often-fatal blood disorder.

Amgen's move allowed it to design oligonucleotide probes that were suitable for genomic screening, one of several methods of cloning. The company was granted a US patent on the gene in 1987 and a European patent in 1990.

Opponents argue that the knowledge and technology for cloning the gene were all openly available. Simply gaining access to the source of the protein does not constitute an inventive step, they argue, a requirement for a European patent.

Amgen dismisses this by saying that its competitors had simply not thought of using this source of erythropoietin, and that its inventive step was to realize that the source could be acquired and used.

Opponents also claim that Amgen did not disclose sufficient information to allow others to repeat the work "without undue burden", another criterion for a European patent. There were mistakes in the sequence of the described genomic DNA and Amgen did not publish the entire sequence, including regulatory coding regions. Moreover, Amgen described the plasmids that it used incompletely and with some errors, although

the company claims that any mistakes do not significantly effect the outcome of the procedures it describes.

Opponents also say that Amgen should have deposited its clone in a publicly accessible repository at the time of US patent filing. Amgen counterargues that its gene is sufficiently described in its patent.

Amgen and GI both feel their patent entitles them to a monopoly of the erythropoietin market. The two companies have been in litigation since 1991, but Amgen's patent has so far proved the more robust, having survived several challenges on both sides of the Atlantic. Amgen now controls all US supplies of the protein and shares the European market with GI's licensee Boehringer Mannheim.

GI lost its original 1987 US patent in 1991, on the grounds that the defined specific activity of its product could not be reproduced by others on the basis of information the company had disclosed in their patent. It was granted a new patent on essentially the same product, but without numerically defining the specific activity, this summer.

At this point, more legal suits started flying. Amgen took out an injunction against GI to prevent it from selling erythropoietin in the United States. And GI filed suit against Amgen's US marketing partner Orthopharmaceuticals for infringing its new patent. If Amgen's patent is upheld in Europe, Boehringer Mannheim will have to bear heavy fines for its infringement.

**Alison Abbott**

## 'Ig' Nobel awards celebrate irredeemable research

Boston. As the Nobel prize award committees were polishing their press releases in Stockholm, on the other side of the Atlantic 1,200 spectators crammed into the Kresge Auditorium of the Massachusetts Institute of Technology on 6 October for Fourth First Annual Ig Nobel Prize Ceremony.

A total of 10 Ig Nobel Prizes were handed out for "irregular, irredeemable, or irreproducible achievements in science, technology, and other irreproachable disciplines". Marc Abrahams, editor of *The Annals of Improbable Research (AIR)*, published by the MIT Museum, described such achievements as "crying out" for recognition.

Robert Lopez, a veterinarian from Westport, New York, captured the entomology prize for experiments in which he inserted ear mites from cats into his ear and painstakingly observed the results. "Immediately, I heard scratching sounds, then moving sounds, as the mites began to explore my ear canal," Lopez reports.

Lopez survived the ordeal and showed

up to receive his "Ig," generously offering to distribute ear mites among the crowd.

Bob Glasgow, a state senator from Texas, won the chemistry award for sponsoring drug-control legislation which now makes it illegal to purchase beakers, flasks, test tubes and other glassware without a special permit. Tim Mitchell of Corning, Inc. was present to receive the prize and praise the new bill.

"It always starts with a test tube, but suddenly that's not enough," Mitchell noted. "Before you know it, you're lying with a beaker in one hand, a flask in the other, strung out and begging for grant money." But he warned that bans on laboratory equipment can be a slippery slope. "Where will it end? Will pocket protectors be next? Remember, beakers don't hurt people. People hurt people."

The prize ceremony provided a unique opportunity to see a live performance of "The Interpretive Dance of the Electrons", featuring three bonafide Nobel Laureates: Dudley Herschbach, William Lipscomb and Richard Roberts. Partici-

pants were also treated to a procession of 'Ignitaries' that includes NonExtremists for Moderate Change in Finland, the Boston Pig Latin Chorus and the 21st-Century Center for 20th-Century Astrology.

The Heisenberg Certainty Lectures were another irreproducible feature of the ceremony. In a 30-second soliloquy, Lipscomb warned the US Congress that "if your position is everywhere, your momentum is zero". Astrophysicist Margaret Geller explained that "everything you've heard about the universe is false". Marvin Minsky, the guru of artificial intelligence, was heckled offstage owing to the dearth of humour in his deliberately ponderous presentation.

In the midst of this excitement, it was easy to forget the runners-up — improbable researchers who didn't quite make the grade. For these Ig hopefuls, Abrahams offered some encouragement: "If you can't help yourself, by all means continue your efforts. If your work is truly irredeemable, one day its worth will be fully appreciated." Steve Nadis



# Overregulated biotechnology

SIR — Nowhere is the gulf between politicians' rhetoric and their actions more pronounced than in US science and technology policy. One need only contrast the negative impact of the administration's biotechnology policies with the lofty goals outlined in the document *Science and the National Interest* (see *Nature* 370, 317; 1994).

Thus the policy document offers as a goal "a stable science-based regulatory system". The reality is that at the same time as funding is down, biotechnology regulation has become increasingly intrusive, unscientific and focused on negligible-risk activities.

The Department of Agriculture (USDA) has required unnecessary permits for more than 1,400 field trials of genetically modified plants, all of them of negligible risk. The Food and Drug Administration (FDA), which has a generally positive 15-year track record in regulating biotechnology, recently announced that it will soon require food manufacturers to notify the FDA before marketing foods manufactured with high-precision recombinant DNA techniques — while exempting those developed with other techniques, regardless of possible risk.

Biotechnology regulatory policies of the Environmental Protection Agency (EPA) are the most egregious of all. In early September, the EPA published regulations for biotechnology control agents. EPA targets only products made with the most precise and predictable new genetic methodologies. The agency also plans to expand its regulatory dominion to a whole new category of products — plants that are made resistant to pests by using the new genetic techniques. These garden and farm plants will be regulated even more stringently than chemicals similar to DDT or parathion. Yet plant breeders have been creating and farmers using genetically improved plant varieties safely for more than a century without government regulation.

These regulations make neither scientific nor economic sense. The USDA, FDA and EPA regulatory approaches fly in the face of a broad scientific consensus that the new biotechnology is an extension, or refinement, of earlier techniques of genetic manipulation. Moreover, their regulatory policies constitute, in effect, a tax on innovation that uses the new biotechnology. Ineluctably, these anti-innovative policies will discourage research using the newer, more precise techniques, denying consumers an array of new products. There is a peculiar irony in the administration's inconsistency: its policies have preferentially undermined research on precisely the kinds of low-value-added but socially important pro-

ducts that should appeal to the administration's policymakers — safer and more nutritious foods, improved bioremediation agents, alternatives to chemical pesticides and fertilizers and other environmentally friendly innovations.

**Henry I. Miller**

Hoover Institution &  
Institute for International Studies,  
Stanford University,  
Stanford, California 94305-6010, USA

## US health reform

SIR — You say that the health plan proposed by a "well-meaning" Clinton administration violates "the Hippocratic oath to do no harm" (*Nature* 371, 2; 1994). If you were ever to read the oath, you would quickly discover that it enjoins the physician to act for the benefit of his patients rather than for "their hurt or for any wrong". In other words, what the oath prohibits is the intentional infliction of damage on patients, not unintended bad outcomes. If Hippocrates had prohibited bad outcomes resulting from treatment by "well-meaning" physicians, no one in ancient Greece would have been qualified to practise medicine.

Politicians in the United States who qualify their positions on health-care reform with similarly twisted renditions of Hippocrates neglect to consider other public-policy consequences of the oath. If Hippocrates were in charge today, there would be no medical schools, because modern doctors are disinclined to work for nothing ("I will teach . . . this art without fee or covenant"), and there would be no health insurance, public or private, because doctors would be prohibited from communicating to third-party payers any information about the identity, condition, or treatment of patients ("Whatever things I see or hear . . . in my attendance on the sick . . . I will keep silence thereon, counting such things to be as sacred secrets").

Elsewhere in the same issue, you publish a communication in which chain letters are classified as "mind viruses". Politically motivated misquotations may not qualify as mind viruses, but they are at least mind bacteria.

**Gareth Penn**

13-1/2 Linda Avenue,  
San Rafael, California 94903, USA

## Origin of life

SIR — In his recent article<sup>1</sup> on the origin of the cell membrane, John Maddox focuses on the model of Ourisson and Nakatani. The features of this model have striking similarities to another theory, not men-

tioned by Maddox, namely that of Günter Wächtershäuser<sup>2</sup>. This latter theory entails the origin of the cell membrane and is based upon an autocatalytic surface metabolism. Wächtershäuser's theory is steadily gaining prominence and has been commented upon by Karl Popper<sup>3,4</sup> and Russell, Hall and Gaze<sup>5</sup>.

A key role in the lipid synthesis section of Wächtershäuser's theory is also played by terpenoids (phosphorylated isoprenoids) bound to a solid surface (pyrite) which spring from an ancient pathway to form lipid constituents which undergo condensations and contribute to growth of the lipid sheet. In a mechanism similar to that of Ourisson and Nakatani described by Maddox, at some point the lipid sheet detaches from the surface to form an enclosure. In both models, synthetic pathways for structural and functional components are included in these cell precursors.

Although the two theories have much common ground with respect to the origin of membranes, not only is Wächtershäuser's theory not referred to by Maddox, but Wächtershäuser does not mention the work of Ourisson (or Nakatani) in either of his major articles<sup>2</sup>.

**Russell Hill**

**Peter Århem**

Nobel Institute,  
Karolinska Institute,  
S-171 77 Stockholm, Sweden

**B. I. B. Lindahl**

Department of Philosophy,  
University of Stockholm,  
S-106 91 Stockholm, Sweden

1. Maddox, J. *Nature* 371, 101 (1994).
2. Wächtershäuser, G. *Microb. Rev.* 52, 482–484 (1988); *Prog. Biophys. molec. Biol.* 58, 85–201 (1992).
3. Popper, K. R. *Nature* 344, 387 (1990).
4. Popper, K. R. *et al. theor. Med.* 14, 167–180 (1993).
5. Russell, M. J. *et al. Nature* 344, 387 (1988).

## Pi remembered

SIR — As a postscript to the silly season, and while  $\pi$  is still in the news (*Nature* 370, 323; 1994), your readers may appreciate a mnemonic for 30 digits of  $\pi$ . The number of letters in each word represents the value of the digit:-

*Que j'aime à faire apprendre un nombre utile aux sages!*

*Immortel Archimède, artiste ingénieur,*

*Qui de ton jugement peut priser la valeur?*

*Pour moi, ton problème eut de pareils avantages*

My source (G. F. Hull, *An Elementary Survey of Modern Physics*, Macmillan, New York; 1938) gives no indication of who wrote this interesting, if not useful, poem, how he derived the value of  $\pi$ , or whether the Editor of *Nature* would have approved of his method of derivation.

**W. E. Ormerod**

The Old Rectory,  
Padworth,  
Reading, RG7 4JD, UK

# Plagiarism or protecting public health?

Erik Millstone, Eric Brunner and Ian White

**Is milk produced using recombinant bovine growth hormone hazardous to animal health? Despite wide publicity, the controversy remains unresolved. Rapid publication of all available data is essential if progress is to be made.**

THE genetically engineered hormone bovine somatotropin (BST; also called recombinant bovine growth hormone) produces an increase in dairy cows' milk yield of approximately 15% (ref. 1). There has been much debate about the effects of BST on the health of cows and humans. In March this year, the Canadian parliament's Agriculture Committee held a hearing on the proposed use of BST in Canada during the course of which Dr Michael Hansen of the US Consumer Policy Institute suggested that Monsanto had blocked publication in the scientific literature of a paper of ours, *BST and animal health: the effect of bovine somatotropin on somatic cell counts* (manuscript available from E.B.). In reply, Dr Robert Collier of Monsanto said of E.M. and E.B. "When you take someone else's data and you submit it without putting their names on it . . . it's called plagiarism"<sup>2</sup>. We deny this allegation.

A recent estimate suggests that Monsanto and several other firms, including Eli Lilly, Upjohn and Cyanamid, have spent approximately £700 million on the development of BST<sup>3</sup>. These companies have been trying for more than 10 years to persuade governments and international regulatory agencies to approve the entry of their products onto the market. In November 1993, the US Food and Drug Administration gave its approval for the sale of Monsanto's BST product (trade name Posilac), but in the European Union BST continues to be subject to a moratorium.

We have been investigating whether the administration of BST to lactating cows increases the number of somatic cells (leukocytes) in their milk. The somatic cell count (SCC) is an indicator of the level of inflammatory response of the udder; an elevated count is associated with an increased risk of mastitis.

In the spring of 1989, researchers at Monsanto published estimates<sup>4,5</sup> of average SCC levels for both treated and control groups from eight randomized controlled trials conducted in the late 1980s, but details were relatively sparse. They observed "a significant increase in somatic cells at 2 sites and in other trials there was no difference in somatic cells"<sup>4</sup>. Significant differences were reported from studies in the United Kingdom and at Cornell<sup>5</sup>.

In the summer of 1989, an undergraduate called Hugh Dawkes, at Chester Col-

lege, submitted a finals project on BST for which E.M. was the external examiner. Dawkes had astutely realized that, although each of Monsanto's eight studies might not, separately, be sufficiently powerful to detect the effect of BST, by pooling the data from all the studies it should be possible greatly to increase the sensitivity of a statistical analysis.

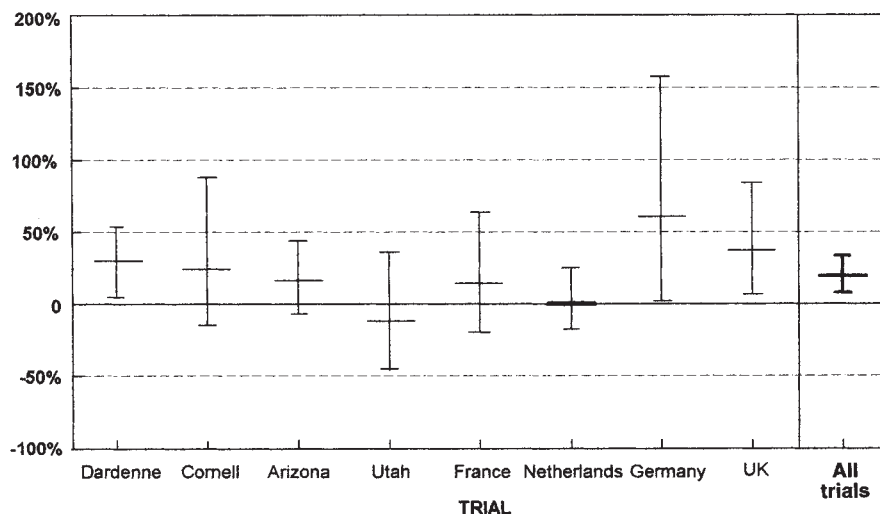
To pursue Dawkes' idea, it was clear that a pooled analysis required supplementary information: at the very least, data on the size of the eight pairs of BST-treated and placebo groups, and the standard deviation of the SCC measurements for each of those groups. E.M. requested this information from Monsanto in August 1989, who promptly provided a table of figures. We then incorporated the new information into a paper submitted to the Veterinary Products Committee (VPC) at the Ministry of Agriculture Fisheries and Food in September 1989 (ref. 6). The committee was then considering the application for a commercial licence for BST, and we had concluded, from the available data, that there was evidence that the milk from cows treated with BST contains statistically significantly increased levels of somatic cells (or, more prosaically, pus).

E.M. then noticed that some of Monsanto's published figures did not coincide with those provided directly to us. A Monsanto scientist, Dr Neil Craven (then manager of the company's Animal Sciences Division), explained that, while it was

the second set of figures that had been submitted to the UK government, neither set was correct owing to mistakes in the arithmetic. Craven and Gerard de Kerchove (Monsanto's senior European biometrician) visited E.M. on 4 October 1989, providing a revised, definitive set of figures and a disk containing the raw SCC data in a spreadsheet file. Two weeks later, Craven wrote to explain why 10 cows that had started the trial were omitted from their eventual analyses. He added: ". . . we request that the raw data be kept confidential. We hope that you will discuss any interpretation of the data with us before disclosing it to third parties".

We decided to carry out a detailed analysis of Monsanto's data. The data provided to E.M. by Craven and de Kerchove were, for each of the 620 cows, the lactation number (age of the cow according to number of pregnancies), weekly SCCs, trial site and trial group (control or BST treatment). Missing values in the SCC data were more frequent towards the end of the cycles of lactation, and in the United Kingdom trial cell counts seemed to have been conducted on a monthly basis. We replicated what Monsanto had indicated were their authentic results, as well as our own previous analysis submitted to the VPC.

Our new analysis incorporated two improvements. First, we used SCC averages over all 7 weeks for the baseline period and all 43 weeks for the trial period, rather



**Effect of BST on mean SCC in Monsanto's trials. Excess percentage increase in treated groups versus controls. A significant effect is evident in those trials where the 95% confidence interval does not cross the zero line (see over for further details).**



than just 2 and 28 weeks as used by Monsanto. Second, we pooled the data from all the trials, thus increasing the effective sample size. Both improvements increased the study's power to detect a change in mean SCCs. For such a meta-analysis to be valid, the studies must be sufficiently similar in three key respects: equal composition of the compared groups in their baseline state; equal performance of the principal treatment; and equality in the detection of outcome<sup>7</sup>. In this case, the eight studies met those conditions.

Our main findings are shown in the figure. The right-hand column shows the pooled analysis, and the vertical bar represents a 95% confidence interval for the effect of BST. This shows that, on average, BST treatment produced a 19% increase in milk SCCs relative to controls; this effect is highly statistically significant ( $P = 0.004$ ). For comparison, the corresponding results for the eight separate studies are shown to the left of the pooled analysis. A higher proportion of cows (10.9 vs 6.0%) fell into the highest EU cell-count band in the BST-treated group, an important criterion because penalty payments are levied on farmers if specific SCC levels are exceeded<sup>8</sup>.

We recognize that our analysis is imperfect in one particular respect. Because SCC and mastitis tend to be related to milk yield — the higher the yield the higher the risk of mastitis — we have not been able to distinguish between the effect of BST *per se* and the effect of increased milk yield resulting from BST use. Until data on milk yields become available such an analysis is not possible. Another issue is whether there were other trials of which we were unaware (see below).

### Publication hurdles

We submitted the paper reporting our findings to the *Veterinary Record*, the journal of the British Veterinary Association, in July 1991. Our paper passed the journal's peer-review process, apart from three small textual amendments, but the editor would publish our paper only with the written consent of Monsanto. We found that Craven had by then left Monsanto, and our request for consent was referred to Dr Doug Hard at Monsanto's Brussels office. In December 1991, Hard refused to agree to the publication of our paper. His contention was that it was for the principal investigators and their associated institutions to publish the results, and he indicated that those parties were preparing papers for publication, even though more than 2 years had passed since Craven had provided the data.

Naturally, we accept that Monsanto and their principal investigators are entitled to decide when and how their raw data are published. But we do not agree that they have any right to prevent or delay the

publication of our work analysing their data. Both Peel and Phipps had already published some analyses of those data, although not in peer-reviewed journals<sup>4,5</sup>. They had volunteered to provide the raw data, and might reasonably expect that other, independent scientists would want to analyse and discuss those data in the peer-reviewed literature. E.M. therefore wrote to Hard arguing that publication of our analysis would not prejudice the publication of their results or analyses, but that on the contrary it would stimulate greater interest in their work. Hard responded in February 1992 by reiterating his refusal to consent to the publication of our paper, and argued that "As the raw data are confidential all subsequent analyses are as well." He added that "We have recently submitted a paper on this subject to the *Journal of Dairy Science* [JDS] and have another in preparation. I will be happy to provide those to you as soon as they are accepted for publication."

At that stage our main concern was to ensure that the attention of the regulatory authorities was drawn to the possible effect of BST on the health of cows and the quality of their milk. We therefore sent copies of our paper to the UK VPC, the EC's Committee on Veterinary and Medicinal Products, the Joint Expert Committee on Food Additives of the World Health Organisation and the UN Food and Agriculture Organisation, and to the US Food and Drug Administration. We marked those papers "not to be quoted without the authors' permission".

By February 1993 we had not received any news of Monsanto's papers, so we suggested to the editor of JDS that Monsanto's paper(s) and ours should be published side-by-side (subject, of course, to normal peer review), thereby stimulating debate. JDS did not invite us to offer our paper for simultaneous publication, and the editor said that "...no papers have been submitted to the *Journal of Dairy Science* by Dr Hard on the topic of bST [sic] and somatic cell counts." A paper from Monsanto on clinical mastitis in cows treated with somatotrope, first submitted in August 1993, however, was published in the summer of 1994 (ref. 10).

Because neither the *Veterinary Record* nor JDS would publish our paper we approached the editor of *The British Food Journal* (BFJ). He said that, if our paper passed their peer-review process, he would not require Monsanto's prior consent because the company had no rights over our analyses and discussion. Our paper was duly passed by the BFJ's peer reviewers, and our work was typeset and scheduled for publication. At that stage the BFJ's commercial publishers required us to indemnify them against breach of copyright. We were not prepared to do this, not because we had any doubts about the authenticity of our copyright, but

because we would have been liable for the consequences of any subsequent decision by the publishers.

Monsanto's legal rights over the raw data are unambiguous, but the issue of rights concerning analyses of their data appears to be a grey area. Specialists on intellectual property rights with whom we have discussed the matter say that we are entitled to publish our analyses, provided that acknowledgement of the sources is made. We have complied with Monsanto's requests to see our work before publication, and to keep the raw data confidential. Whatever the legal position may be, we entirely reject any suggestion that we have plagiarized the work of Monsanto's scientists. We have fully acknowledged the source of the data. Our paper simply seeks to make public results that Monsanto appears to be making little effort to publicize, and that we believe are of importance to the debate over the licensing of BST.

In the course of giving evidence to the Canadian parliament in March, Collier made reference to the forthcoming JDS paper<sup>10</sup>, written by him, Craven, de Kerchove, Hard and 22 others, which discusses clinical mastitis in the BST trials. This paper differs from ours in two key respects. First, it reviews data not from 8, but from 15, studies. Second, the authors do not provide a full analysis and discussion of the SCC data that we have tried to publish. Hard declined to provide us with the data from the seven additional studies. Until those data are in the public domain, some important questions about the effects of BST on animal health will remain unresolved. □

Erik Millstone is at the Science Policy Research Unit, Sussex University, Brighton BN1 9RF; Eric Brunner is at the Department of Epidemiology & Public Health, University College London, London WC1E 6EA; and Ian White is at the Medical Statistics Unit, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK.

1. *Bovine Somatotropin — An Introductory Guide to a New Farming Tool* (Monsanto, Brussels, November 1986).
2. *Minutes of the Standing Committee of Agriculture and AgriFood* p. 7 (Canadian House of Commons, Ottawa, 10 March 1994).
3. Morse, L. & Maitland, A. *Financial Times* p. 22 (2 March 1994).
4. Peel, C. J., Eppard, P. J. & Hard, D. L. in *Biotechnology in Growth Regulation* (eds Heap, R. B. et al.) p. 110 (Butterworths, London, 1989).
5. Phipps, R. H. in *The Use of Somatotropin in Livestock Production* (eds Seijrsen, K. et al.) 88 (Elsevier, London, 1989).
6. Millstone, E. P., Barrell, R. J. & Brunner, E. J. *BST and Mastitis: a preliminary statistical analysis of available data* (submitted to the Veterinary Products Committee, London, September 1989).
7. Gerbarg, Z. B. & Horwitz, R. J. *clin. Epidem.* **41**, 503–509 (1988).
8. *Milk Marketing Board, Mastitis and Cell Counts* (Milk Marketing Board Communications Division, Thames Ditton, 1993).
9. Ahmad, S. R. & Choo, V. *The Lancet* **341**, 1018–1019 (1993).
10. White, T. C. et al. *J. Dairy Sci.* **77**, 2249–2260 (1994).

# Star masses and bayesian probability

**The use of Bayes' theorem to pin down the mass limits of the few neutron stars found in binary systems is both a splendid illustration of the explicit use of *a priori* knowledge and a means to a useful result.**

THE general belief that bayesian probability is simply a way of making bricks without straw (or inferences without data) is well entrenched. The late Sir Harold Jeffreys, one of the exponents of the view that, in estimating the probability of some event, one should be free to take account of prior information, even of a subjective character, gives a neat counter-example in the preface of his book on the subject. He supposes that the reader visits an unfamiliar city (Birmingham in England, as it happens), catches a bus to make a short journey and observes that the bus carries a serial number (say  $N$ ) on the glass partition just behind the driver. So how many buses are there in the city?

The orthodox position is that it is impossible to tell. There is only one datum, the number  $N$ . Even if one is told that the city's buses are numbered consecutively, from 1 to some maximum, the problem is strictly indeterminate. Yet, says Jeffreys, the incautious or the statistically naive will give the answer  $2N$ . And, the argument goes, there is a sense in which the answer, while almost certainly incorrect, is more useful than no answer at all. At least it makes some use of the information gleaned from the glass behind the driver's head as well as of the general knowledge that, however large the city, the number of buses will not be infinite.

There should therefore be a measure of compassion for an attempt by Lee Samuel Finn to pin down the probability distribution of the masses of neutron stars on the basis of just four recent observations, all of them binary pulsar systems in which the (radio) silent component is believed also to be a neutron star. But Finn's argument is not naive guesswork in which a *a priori* probability is of the essence, but rather a sober use of Bayes' uncontroversial theorem on conditional probabilities where the *a priori* knowledge is used only in the most explicit way. It turns out to be a neat illustration of how to make bricks out of very little straw.

The two best-determined pulsar systems in which both components are neutron stars are PSR 1913+16 and PSR 1534+12, first described by J. H. Taylor and J. M. Weisberg (*Astrophys. J.* **345**, 434; 1989) and A. Wolszczan (*Nature* **350**, 688; 1991) respectively. So far, these are the only binary pulsars in which both components are neutron stars and in which it has been possible to determine the masses with a high degree of precision. Finn notes (*Phys. Rev. Lett.* **73**, 1878-1881; 1994) that the masses are all remarkably alike. In the two systems, the total mass amounts to  $2.828M_{\odot}$  and  $2.679M_{\odot}$ ,

where  $M_{\odot}$  is the solar mass. Similarly, the masses of the silent companions are estimated at  $1.442M_{\odot}$  and  $1.36M_{\odot}$  respectively.

What then to make of these data? On the assumption that there is no intrinsic difference between the silent and the pulsating stars, there are four numbers to play with, so that it is possible to calculate a mean and a standard deviation. The arithmetic, for what it is worth, is that the mean mass of a neutron star is  $1.376 \pm 0.031M_{\odot}$ .

The trouble with that result is that it imparts very little confidence. A single discordant measurement from elsewhere, say  $2.00M_{\odot}$ , would yield a much greater value of the average mass and a very much greater value of the standard deviation. So why not calculate confidence limits for the estimate of the mass of a neutron star? To do that in a meaningful way, it is necessary to make assumptions of some kind about the form of the distributions of the masses of neutron stars, which may in itself be hazardous.

In reality, there are a few more data to play with. One such system is the binary pulsar PSR 2303+46, where both objects are neutron stars but where the observations are not yet sufficient to disentangle from the estimation of the masses of the two companions quantities such as the inclination of the orbit and the line of sight or, what comes to the same thing, the semi-major axis of the mutual orbit. The other binary system in this class is PSR 2127+11C.

Again it is necessary to assume a distribution for the masses of neutron stars, which Finn assumes to be uniform between lower and upper limits, say  $m_l$  and  $m_u$  respectively. Then, clumsily in words (with the conditional restrictions in parentheses), the probability that the limits are indeed these (given the data and the distribution law) is equal to the probability of the data (given the supposed limits and the *a priori* knowledge) multiplied by the probability of the mass limits (given the *a priori* knowledge) and divided by the probability of the data (given the *a priori* knowledge).

What *a priori* knowledge is there? It seems to be agreed that neutron stars would become black holes if they were more massive than  $3M_{\odot}$ , while the equation of state of neutron matter suggests that a neutron star could not be less massive than about  $0.1M_{\odot}$ . These are very wide limits which do not rely at all on the now classical Chandrasekhar limit for the size of the degenerate core of a star from which hydrogen has been exhausted, and which would fix the mass of a neutron star at something less than  $1.4M_{\odot}$ .

The outcome of the algebra is straightforward, but with a few pitfalls. For example, the probability of the limits  $m_l$  and  $m_u$  (given *a priori* knowledge) requires that each should be uniformly distributed in the allowed range and that the upper limit should always be greater than the lower. The virtue of the calculation is that everything is explicit. In particular, the formalism makes it plain how information from extra data may be simply thrown into the pot as it accumulates. For what it is worth, the data from four pulsars yield (with 95 per cent confidence) a lower bound between  $1.01$  and  $1.34M_{\odot}$  and an upper bound between  $1.43$  and  $1.64M_{\odot}$ .

It is important that these numbers are estimates of upper and lower extremes of a uniform probability distribution for the masses of real neutron stars. It will be interesting to see how quickly the limits close up upon each other as further data accumulate. Meanwhile, it seems inevitable that this example will quickly find its way into some textbook as an illustration of how inferences can be drawn from a meagre collection of data. In this case, of course, students will not be tempted away from the path of political correctness by the use of *a priori* knowledge which, being explicit and objective, is unexceptionable.

Finn's purpose, though, is more immediate than pedagogical. He appears to be one of the growing army of people who are awaiting the time when the gravitational wave detector called LIGO bursts on the world some years from now. The point is that neutron-star binaries should be sources of gravitational radiation that will, when the parameters are right, be detectable at LIGO and other such instruments. Winning a feel for the distribution of neutron star masses in the real world is a necessary first step towards telling when LIGO signals will be meaningful.

Indeed, that nicely illustrates the problem of drawing inferences from uncomfortably few data. At some future stage, when LIGO has been commissioned, it will be necessary to look for gravitational signals with a period corresponding to the orbital period of the binary, and to estimate from that a quantity which is a function of the product and the sum of the two stellar masses. As Finn puts it: "An accurate assessment of our prior knowledge is especially important in determining when a signal is sufficiently strong that it refines our understanding as opposed to affirming our existing prejudices". There could hardly be a more apposite defence of Bayesian probability.

John Maddox



# Folding a jelly sandwich

Maria T. Zuber

WITH an average elevation of 5 kilometres and a surface area of over 5 million square kilometres, the Tibetan plateau is the largest topographic upland on Earth. Its association with the actively rising Himalayan mountains dramatically illustrates the forces associated with the global motions of the Earth's surface plates. But despite its prominence, key characteristics

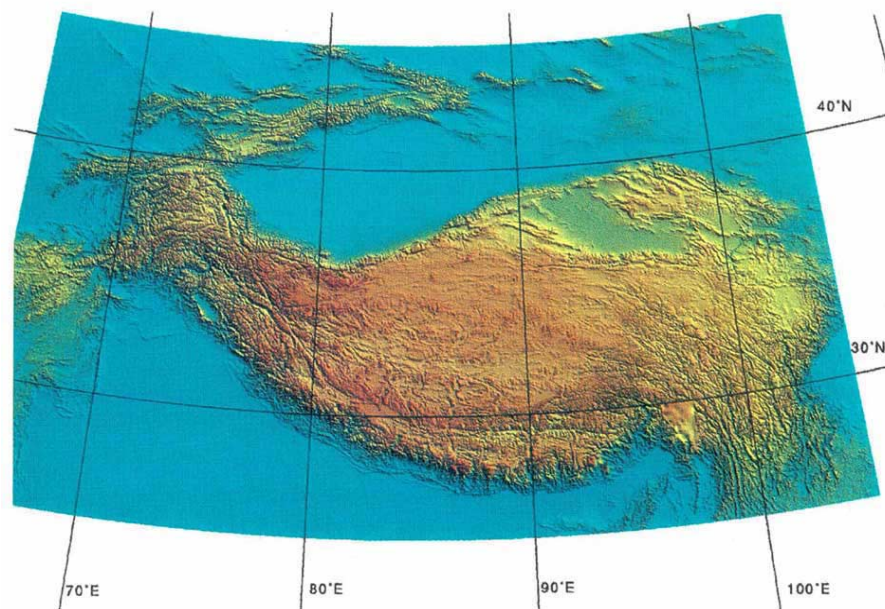
depths and both the crust and mantle must be weak. In another model, the crust of India was injected into a weak region in the Asian lower crust<sup>3</sup>. Here the surface of the plateau would be largely undeformed but deformation would be extensive in the lower crust, and the upper mantle might be subducting. A third possibility is that Asian crust was thrust over India<sup>4</sup>. In this

face topography and of subsurface masses that deviate from the 'average' density structure. Such deviations are usually attributable to either thermal or compositional differences. By subtracting the attraction of surface topography from the gravity signal, it is possible to map out subsurface density variations. In their analysis, Jin and colleagues identified two wavelengths in the gravity of Tibet which they interpret as undulations in subsurface density interfaces, one of which corresponds to the boundary between the crust and mantle.

Their analysis suggests that the plateau is folding at two distinct wavelengths, 150 and 500 km. The east-west orientation of the folds is consistent with the source of compressive forces being the India-Asia collision. Theoretical models of collisional deformation have shown that two wavelengths of folding can be produced if the shortening medium contains two strong layers separated by a weak ductile layer<sup>7</sup>. With this in mind the data are most simply explained by a 'jelly sandwich' model of continental strength, in which the upper crust and upper mantle are strong and are separated by a weak, ductile lower crust.

The distribution of strength with depth is governed mainly by the compositions of the crust and mantle coupled with the temperature structure of the medium. Preliminary analyses of relationships between gravity and topography indicate that the observed wavelengths are consistent with the proposed strong and weak layer thicknesses for a continental crust of twice the normal thickness. The gravity data further show that the upper mantle of Tibet seems to be too strong to shorten by the amount needed to double the thickness of the crust. This has led Jin and colleagues to favour a model in which a strong upper mantle subducts under the plateau, eliminating models for the formation of the plateau that require a weak mantle.

Their results provide an enticing glimpse beneath the surface of Tibet, but many questions about the structure and tectonics of this region remain. It will be important to compare the gravity and topography data to compressional folding models, and to analyse this data further to substantiate or disprove the possibility of subduction. The evolution of deformation



Folding up? Shaded relief image of the Tibetan plateau<sup>8</sup>, the largest upland region on Earth. Illumination is from the northeast. Blue is low and red high. The dynamic range of topography in the figure is 7 kilometres. Image courtesy of Eric Fielding, Cornell University and Jet Propulsion Laboratory.

such as the subsurface structure of the plateau, the manner and extent to which it is deforming, and its influence on global tectonics are not understood. A central question in understanding the nature of tectonic deformation is the vertical distribution of strength in the crust and mantle. On page 669 of this issue, Jin and others<sup>1</sup> use gravity data to reveal evidence for large-scale folding of the Tibetan crust that implies the presence of a weak lower crustal layer sandwiched between a strong upper crust and upper mantle.

The plateau is the manifestation of the collision of India with Eurasia that started 45 million years ago, and resulted in a doubling of the thickness of the crust. Previous suggestions to explain the crustal thickness in the context of collisional deformation fall into three classes, each of which makes specific predictions about how deformation should be distributed in the crust and upper mantle.

In one of these, India acted as a 'rigid indenter' into Asia<sup>2</sup>, resulting in a piling up of Asian crust. In this case the plateau would be considerably deformed at all

case there should be significant deformation at the surface and possibly subduction of the Asian mantle beneath the plateau<sup>5</sup>. However, the sparse geological and geophysical information that has been available for Tibet has been inadequate to distinguish between these models.

Until now, folding of Tibet's crust has not been recognized, although folding is a common consequence of continental collisions. In fact, 200-km-wavelength folds on the sea floor southeast of India are attributed to the India-Asia collision<sup>6</sup>. The absence of large-scale folds in Tibet would indicate that crust is too weak to fold and instead just thickens, as folding requires a high-strength layer which bends coherently in response to compressive forces. The thickness of the strong layer, which provides key information on the thermo-mechanical structure of the compressed medium, can be deduced from the wavelength of folding. For Tibet, Jin and colleagues have now found evidence for folding in previously unreleased gravity data.

Gravity measures the attraction of sur-

1. Jin, Y., McNutt, M. K. & Zhu, Y. *Nature* **371**, 669–674 (1994).
2. England, P. & McKenzie, D. P. *Geophys. J. R. astr. Soc.* **70**, 295–321 (1982).
3. Zhao, W.-L. & Morgan, W. J. *Tectonics* **4**, 359–369 (1985).
4. Powell, C. M. & Conaghan, P. J. *Earth planet. Sci. Lett.* **20**, 1–12 (1973).
5. Willett, S. D. & Beaumont, C. *Nature* **369**, 642–645 (1994).
6. Weissel, J. K., Anderson, R. N. & Geller, C. A. *Nature* **287**, 284–291 (1980).
7. Zuber, M. T. *J. geophys. Res.* **92**, 541–551 (1987).
8. Fielding, E., Isacks, B., Barazangi, M. & Duncan, C. *Geology* **22**, 163–167 (1994).

must also be considered: how do the results of the present study, which imply significant near-surface deformation, tie in with observations of plateau topography that have been interpreted as evidence of very little deformation of the upper crust in recent times<sup>8</sup>? And what is the role of erosion and post-collisional effects such as gravitational spreading of plateau topography?

Finally, we should ask ourselves whether the actively deforming Tibet–Himalayan deformation zone is unique, so that we can evaluate the implications of these new observations for the general

process of how continents collide. For example, the Andean collision zone in South America has produced an adjacent plateau, the Altiplano, whose structure and origin is also a matter of debate. Does the Altiplano, too, have a jelly sandwich strength structure and should crustal-scale folding be expected there as well? □

Maria T. Zuber is in the Department of Earth and Planetary Sciences, Johns Hopkins University, Baltimore, Maryland 21218, USA, and the Geodynamics Branch, NASA/Goddard Space Flight Center, Greenbelt, Maryland 20771, USA.

## CHEMICAL BONDS

# A topological classification

David L. Cooper

WHAT determines the form of a chemical bond? Many of the qualitative concepts currently used to describe molecular electronic structure emerged in the earliest days of quantum mechanics. Indeed, one of the first successful attempts to account for chemical bonding was published by Gilbert Lewis in 1916, several years before the advent of the Schrödinger equation. Lewis structures, which emphasize the importance of the shared electron pair, are still in widespread use. But a more rigorous approach, firmly founded in quantum mechanics, is highly desirable, especially if it can be made quantitative.

Silvi and Savin describe on page 683 of this issue<sup>1</sup> a new classification of chemical bonding which is based on the analysis of a physical observable, namely the electron density, while taking proper account of the influence of Pauli's exclusion principle. Their approach leads to a clear qualitative demarcation between shared-electron interactions, as in covalent and metallic bonding, and unshared-electron interactions, as in ionic and hydrogen-bonded systems.

It was the pioneering work in 1928 of the late Linus Pauling that first showed how directed covalent bonding, based on the pairing of electrons, could be rationalized by theoretical quantum chemistry. Pauling's book *The Chemical Bond* (1948)<sup>2</sup> and Coulson's *Valence* (1961)<sup>3</sup> have both had a profound influence on the development of our understanding of molecular

electronic structure. Many of the concepts described therein have largely been vindicated in recent years by respectable quantum chemical calculations, although some have been modified and others abandoned. New concepts have also emerged. Advances are being made using various *ab initio* techniques, such as modern forms of valence-bond theory, which combine ease of interpretation with accurate numerical results, and by analysis of density matrices. Of course, such methods typically rely on particular approximations to the wavefunction, whereas it could be much more appealing to obtain information

directly from physical observables such as the electron density,  $\rho(\mathbf{r})$ .

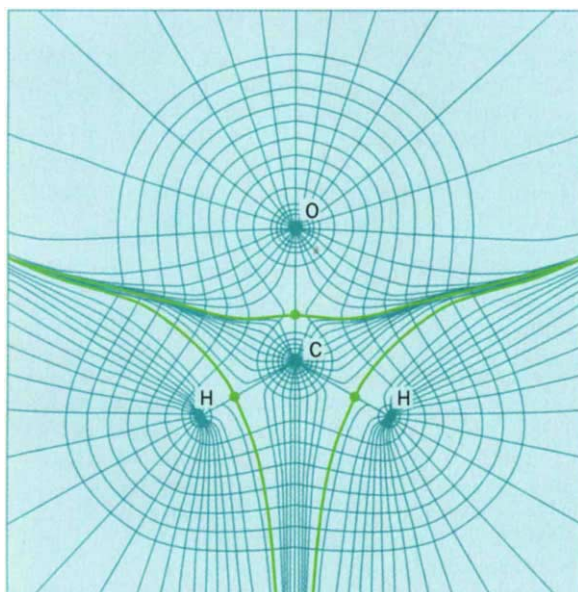
Probably the most widely used direct analysis of  $\rho(\mathbf{r})$  is the one developed over many years by Richard Bader<sup>4</sup>, in which the electron density for a molecule or crystal is split up into densities for smaller fragments. The method involves locating the steepest downhill paths through the electron density. Such paths are called the trajectories of  $\nabla\rho(\mathbf{r})$ , with the term 'trajectory' stemming from the allusion to rolling a marble or ball-bearing across a nonplanar surface. A zero-flux surface through the electron density is one for which there is no net flow of charge perpendicular to the surface. Such surfaces consist of those trajectories of  $\nabla\rho(\mathbf{r})$  that originate at infinity and terminate at the saddle points of the electron density in the internuclear region, as the figure illustrates.

Bader demonstrated that the use of zero-flux surfaces to partition a molecule or crystal leads to equations of motion which mirror exactly those for the total system. The sets of trajectories  $\nabla\rho(\mathbf{r})$  which terminate at the nuclei (attractors) span the space bounded by the zero-flux surfaces and define atomic basins, as shown in the figure. Atomic properties, such as multipole moments, energies and forces, may be obtained in a rigorous and consistent manner by appropriate integration over the basins. Particular emphasis is also placed in Bader's approach on the analysis of the laplacian,  $\nabla^2\rho(\mathbf{r})$ .

An interesting alternative is based on examining the ways in which the shapes or topological types of three-dimensional contours (isosurfaces) of  $\rho(\mathbf{r})$  change as the threshold for the isosurface is varied<sup>5</sup>. This type of scheme has proved especially useful in studies of molecular similarity<sup>6</sup>, addressing the question "How similar are the electron densities of two particular molecules?"

The new approach of Silvi and Savin complements and augments the theory of Bader, while using the language of topological techniques. In essence, Silvi and Savin employ an electron localization function which is a measure of the influence of Pauli's exclusion principle on the local kinetic energy of an assembly of fermions, such as electrons. This is a relatively simple function of  $\rho(\mathbf{r})$  and  $\nabla\rho(\mathbf{r})$ , and takes values between zero and one.

The excess kinetic energy takes low values in regions occupied by paired electrons, and this leads to values of the electron localization function close to unity. Isosurfaces of the electron localization function define localization domains which may contain one or more local maxima, characterized as point, ring and spherical attractors. Bifurcations (or changes of topological type) occur at critical values of the isosurface threshold



Contours of the electron density  $\rho(\mathbf{r})$  in the molecular plane of formaldehyde,  $\text{H}_2\text{CO}$  (adapted from ref. 7). The locations of the saddle points in the electron density are marked with green circles. Two types of trajectory of  $\nabla\rho(\mathbf{r})$  have been superimposed. The few lines that terminate at the saddle points define the zero-flux surfaces and partition the molecule into fragments. The rest of the trajectories, which terminate at the nuclei (attractors), span the spaces bounded by the zero-flux surfaces and define the atomic basins in the approach of Richard Bader.



and reduce one or more localization domains into a larger number of domains containing fewer attractors.

The key consideration in the method of Silvi and Savin is the spatial arrangement and the types of the *irreducible* localization domains. For the molecules that they consider, the locations of the core, bonding and non-bonding attractors appear to be consistent with conventional representations of the electronic structure.

In the absence of symmetry, the Pauli principle implies saturation at two electrons per localization basin. As the authors indicate, an important next step is the integration of  $\rho(r)$  over the basins, so as to provide information on the degree of saturation of the attractors and on the bond multiplicity. Their analysis has been restricted so far to a specific class of wavefunctions for closed-shell systems. It will be important to extend these studies to other spin multiplicities and to excited states, and to analyse electron densities derived from more accurate calculations. Experimentally determined electron densities do not yet achieve the level of accuracy required for a full analysis.

Silvi and Savin have presented us with the first step in the development of an

exciting new approach to the classification of chemical bonding, requiring knowledge only of the electron density. The utility of their work will depend to a large extent on how easily it may be used to compare different geometrical arrangements and different electronic states of families of closely related molecules, and how well it can classify the changes that occur during the course of a chemical reaction. I will be particularly interested to see how it copes with 'difficult' molecules such as FOOF, the fluorinated equivalent of hydrogen peroxide, a deceptively simple-sounding molecule whose bonding nevertheless remains controversial. □

David L. Cooper is in the Department of Chemistry, The University of Liverpool, PO Box 147, Liverpool L69 3BX, UK.

1. Silvi, B. & Savin, A. *Nature* **371**, 683–686 (1994).
2. Pauling, L. *The Nature of the Chemical Bond*, 3rd Edn (Cornell Univ. Press, 1948).
3. Coulson, C. A. *Valence*, 2nd Edn (Clarendon, Oxford, 1961).
4. Bader, R. W. F. *Atoms in Molecules: A Quantum Theory* (Oxford Univ. Press, 1990).
5. Mezey, P. G. In *Reviews in Computational Chemistry*, Vol. 1 (eds Lipkowitz, K. B. & Boyd, D. B.) (VCH, New York, 1990).
6. Duane-Walker, P., Artera, G. A. & Mezey, P. G. *J. comput. Chem.* **12**, 220 (1991).
7. Cooper, D. L. & Stutchburg, N. J. C. *Chem. Phys. Lett.* **120**, 167–172 (1985).

## LONG-TERM POTENTIATION

# A question of reliability

Graham L. Collingridge

SYNAPTIC plasticity is thought to be fundamental for learning and memory in vertebrates. The intense interest in *N*-methyl-D-aspartate (NMDA) receptor-dependent long-term potentiation (LTP) in the hippocampus stems from the belief that it provides a way to understand these synaptic mechanisms (see ref. 1 for review). The most controversial matter exercising students of LTP is whether its expression is pre- or postsynaptic, and in their report on page 704 of this issue<sup>2</sup> Stevens and Wang claim to have pinned down the answer. They conclude that LTP is generated presynaptically. Elegant as their experiments are, however, the case cannot yet be considered closed.

Stevens and Wang applied a small stimulus to activate a single afferent fibre and directly measured the proportion of stimuli that evoked an excitatory postsynaptic current (EPSC) before and following the induction of LTP. The results they obtained were clear-cut — LTP increased the number of successful stimuli but did not affect the size of the EPSC evoked by each successful stimulus. An increase in synaptic reliability (that is, a reduction in failures) has been observed in most previous studies of this sort but has always been accompanied by an increase

in the size of the EPSC (ref. 1). Stevens and Wang conclude that LTP is due to a pure presynaptic change which is caused by an increase in the likelihood of an action potential causing release of neurotransmitter.

What, then, are the caveats? One, as mentioned above, is that other groups have seen changes in measures of the size of the EPSC (in particular, an increase in the separation of the peaks in the histogram of EPSC amplitudes). Unfortunately, Stevens and Wang hardly confront this bone of contention. The inference, however, is that previous studies which used 'minimal' stimulation actually activated multiple fibres. But this explanation cannot, for instance, apply to the study of Malinow<sup>3</sup>, in which LTP was measured between synaptically coupled pairs of neurons.

Another consideration is that the reduction of failures is a form of plasticity which, although it is dependent on the NMDA receptor, is caused by a long-lasting change in presynaptic fibre excitability and is therefore something quite distinct from classical LTP<sup>4</sup>. In theory, a reduction in failures could result from a decrease in the threshold for eliciting the presynaptic action potential or from an

increase in the likelihood of an action potential invading a terminal, as well as (as Stevens and Wang assume) from a genuine modification of the release process.

Finally, there is the possibility that a reduction in failures could result from a postsynaptic change; that is, for example, from the rapid recruitment of functional AMPA ( $\alpha$ -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid) receptors to previously 'silent' synapses. This process could result from AMPA receptor clusters being incorporated into the postsynaptic membrane or by 'silent' AMPA receptor clusters being switched into an active state, perhaps by phosphorylation. Indeed, earlier this year Kullmann<sup>5</sup> presented some evidence that this may occur. Therefore, the results from minimal stimulation studies need to be interpreted in conjunction with findings obtained in other ways<sup>1</sup>.

One method employed has been to measure the release of radiolabelled or endogenous glutamate before and after the induction of LTP. In most studies, an increase has been detected, consistent with a presynaptic change. A second approach has been to follow paired-pulse facilitation; here most reports show no change following the induction of LTP, and the authors concerned have argued for a postsynaptic change. A third tack, taken by myself and colleagues, has been to measure the sensitivity of potentiated neurons to AMPA receptor agonists<sup>6</sup>. An increase in sensitivity, sufficient to account for LTP, was seen but it occurred quite slowly (at a rate coincident with the conversion of short-term potentiation into LTP proper). These results are therefore consistent with both pre- and postsynaptic changes.

Recently, Manabe and Nicoll<sup>7</sup> employed a new approach based on the use of the uncompetitive NMDA antagonist MK-801 to monitor release probability at glutamate synapses<sup>8,9</sup>. The method relies on the use-dependent, slowly reversible, blocking of NMDA channels by MK-801. In the presence of a uniform concentration of MK-801, NMDA receptor-mediated synaptic transmission is progressively blocked at a rate dependent on release probability. Manabe and Nicoll found that this measure of release prob-

1. Bliss, T. V. P. & Collingridge, G. L. *Nature* **361**, 31–39 (1993).
2. Stevens, C. F. & Wang, Y. *Nature* **371**, 704–707 (1994).
3. Malinow, R. *Science* **252**, 722–724 (1991).
4. McNaughton, B. L., Shen, J., Roa, G., Foster, T. C. & Barnes, C. A. *Proc. natn. Acad. Sci. U.S.A.* **91**, 4830–4834 (1994).
5. Kullmann, D. M. *Neuron* **12**, 1111–1120 (1994).
6. Davies, S. N., Lester, R. A. J., Reymann, K. G. & Collingridge, G. L. *Nature* **338**, 500–503 (1989).
7. Manabe, T. & Nicoll, R. A. *Science* **265**, 1888–1892 (1994).
8. Rosenmund, C., Clements, J. D. & Westbrook, G. L. *Science* **262**, 754–757 (1993).
9. Hessler, N. A., Shirke, A. M. & Malinow, R. *Nature* **366**, 569–572 (1993).

ability is unaffected by LTP and argue, therefore, that LTP is postsynaptic in origin. This conclusion agrees, as expected, with the results from paired-pulse facilitation experiments.

This work is also likely to have its critics, however. First, LTP was monitored as a change in the AMPA receptor component of the EPSC rather than of the NMDA receptor component. The latter would have been preferable because it is this which, of course, is used to monitor the MK-801 block. Second, LTP was induced by pairing; this is a procedure which may selectively induce LTP at synapses that already have a high re-

lease probability<sup>8</sup>. Third, the MK-801 was applied at least an hour after LTP had been induced so that the possibility of presynaptic changes at earlier times could not be assessed. Nonetheless, the fact that this measure of presynaptic function provided a negative result, at a time when the postsynaptic AMPA-sensitivity study<sup>6</sup> provided a positive result, points strongly towards postsynaptic modification in LTP.

So can the starkly contrasting conclusions of Stevens and Wang and Manabe and Nicoll be reconciled? One possibility that they can is if, indeed, the increase in reliability is due to rapid recruitment of

AMPA receptor clusters. But Stevens and Wang never saw multiple quanta EPSCs, so there would have to be a mechanism to restrict release to one active zone per action potential (otherwise sometimes simultaneous release from multiple sites would result in multiple quantal EPSCs).

Regrettably, one must conclude that the locus of expression of LTP has not yet been unequivocally identified. LTP is a stubborn problem — on that one can rely. □

Graham L. Collingridge is in the Department of Anatomy, School of Medical Sciences, University of Bristol, Bristol BS8 1TD, UK.

## OBITUARY

## André Lwoff (1902–1994)

ANDRÉ Lwoff, who died on 30 September, was one of the fathers of molecular biology. He began his scientific life at the age of 19, and during the course of his career made a major contribution to the transformation of biology from a collection of dispersed disciplines into a single, unified science. In 1965, together with Jacques Monod and myself, he was awarded the Nobel Prize in Physiology or Medicine for "discoveries concerning the genetic control of enzyme and virus synthesis".

André Lwoff was born to parents of Russian origin; his father was a psychiatrist, and his mother a painter and sculptor. His taste for research was probably influenced by his father who, in the late nineteenth century, had had to escape from oppression by the Tsarist regime like many progressive people of the time. His father believed unshakeably in science, and particularly in Darwinism.

When André Lwoff started in biology, he worked under a great protozoologist, Edouard Chatton, and was fascinated by the beauty and the strangeness of the forms that appeared under the microscope. At the age of 20, he entered the Pasteur Institute, in the laboratory of Felix Mesnil, who had been the secretary of Pasteur himself. During the 1930s, Lwoff spent time in two centres of biochemistry (with Otto Meyerhof in Heidelberg and David Keilin in Cambridge) and in 1938 became head of the Service de Physiologie Microbienne located in the attic of the Pasteur. During the first year of the war, he was in the army. He then came back to the Institute, where his laboratory became an active centre of the Resistance, and only at the end of the war did he resume serious scientific research. In 1959, he was ap-

pointed professor of microbiology at the Sorbonne.

Successively a protozoologist, a bacteriologist, a biochemist, a geneticist and a virologist, Lwoff produced a substantial scientific opus that includes two main discoveries. The first was to do with the status and role of vitamins. To multiply,

found throughout the living world.

The second discovery was the demonstration that the genetic material of a virus, a bacteriophage, can become a component of the genetic make-up of the host bacterium. The whole progeny of this bacterium inherits this property. Yet the equilibrium between the virus and the host cell can be broken by treatments which force the cell to produce virus particles. This notion underlies much of today's work on cancers and retroviruses.

André Lwoff had few students. He liked to work by himself, with the help of his wife Marguerite, and of a technician and one or two collaborators (who, after the war, included Jacques Monod, Elie Wollman, Pierre Schaeffer and myself). In the Pasteur attic, there was an uninterrupted stream of foreign researchers and visiting students, and André generated an exceptional ambience, an atmosphere mixing enthusiasm, clear-mindedness, non-conformism, humour and friendship. Those who had not worked in the attic were unable to perceive the warmth and generosity of this remarkable personality, who shielded himself behind a haughty reserve.

A real Renaissance man, André Lwoff liked to paint, especially in his later years, and he also wrote several books (*L'Évolution physiologique*, *L'Ordre biologique* and *Jeux et Combats*). He functioned more by intuition than by method, and practised science as an art. Indeed, this great scientist was above all an artist.

François Jacob

François Jacob is in the Département de Biologie Moléculaire, Institut Pasteur, 25 rue du Dr Roux, 75724 Paris Cedex 15, France.

IMAGE  
UNAVAILABLE  
FOR COPYRIGHT  
REASONS

Lwoff — a real Renaissance man.

some microbes require traces of certain compounds which other microbes do not need. Before the war, Lwoff showed that these compounds — vitamins — are actually components of all living organisms. They are indispensable to life. If some organisms but not others require vitamins in their culture medium, it is because the latter but not the former are able to manufacture them. This was a far-ranging discovery — with the results stemming from the work of biochemists, it showed that the same molecules with the same functions are



# Incisions for excision

Jan H. J. Hoeijmakers and Dirk Bootsma

EXCISION implies two incisions. Plumbers and surgeons know this principle, and long ago it was put into practice by the evolutionarily ancient DNA-repair machinery. Dissection of the first incision made by the eukaryotic nucleotide-excision repair pathway has now been described by O'Donovan *et al.*<sup>1</sup>, and dissection of the second by Bardwell *et al.*<sup>2</sup>. When put together the two processes enable the replacement of a damaged piece of DNA by a new one.

Genetic information is continuously subject to the deleterious effects of environmental agents. The ultraviolet component of sunlight, for instance, promotes dimerization of flanking pyrimidines, which obstructs transcription and upon replication can lead to permanent mutations. The consequences are obvious. Cells may die or (perhaps even worse) suffer from malfunctioning, with carcinogenesis or inborn errors as a final outcome. These consequences are apparent from a class of cancer-prone genetic instability syndromes which stem from a defect in one of the repair systems that nature has acquired to cope with DNA injury.

An example is xeroderma pigmentosum<sup>3</sup>, which manifests itself as acute sun sensitivity and other skin abnormalities, notably a strong predisposition to skin cancer. Normally sun-induced DNA lesions are removed by nucleotide-excision repair, a broad-spectrum DNA-repair system. Its biochemical complexity, which was already apparent from the large number of genetic complementation groups identified within xeroderma pigmentosum, gained an extra dimension with the discovery of two related disorders, Cockayne's syndrome and trichothiodystrophy (otherwise known as brittle hair syndrome). Both of these neurodevelopmental diseases again feature genetic heterogeneity. Thus there are seven xeroderma pigmentosum groups (XP-A to XP-G), two Cockayne's syndrome groups (CS-A and CS-B), three combined XP/CS groups (coinciding with XP-B, XP-D and XP-G) and three trichothiodystrophy groups (TTD-A, XP-B and XP-D), making a total of at least ten

distinct genes involved in nucleotide-excision repair (see our News and Views article of May last year<sup>4</sup>).

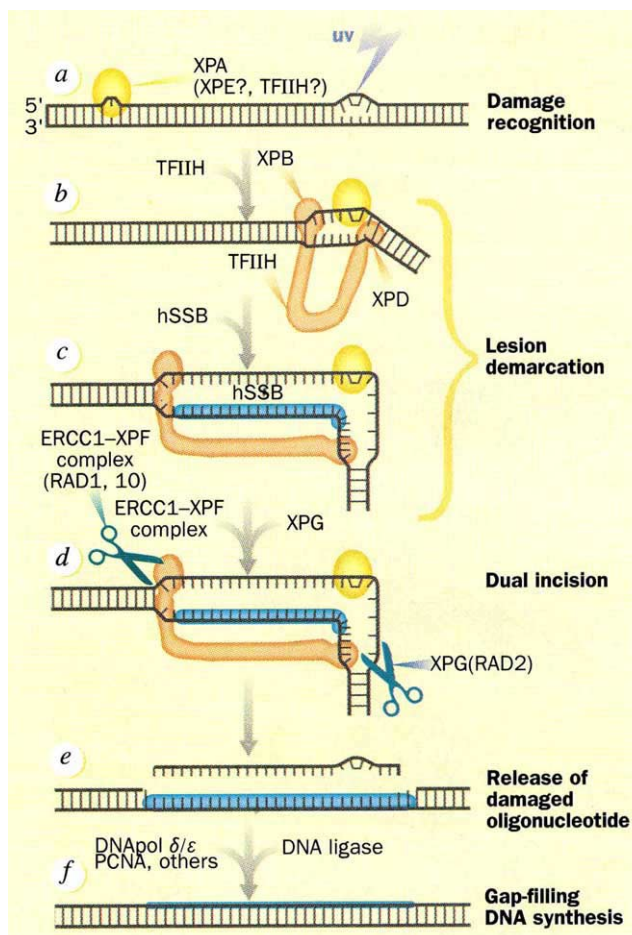
Using synthetic 'splayed-leg' and 'bubble' DNA substrates, O'Donovan and colleagues<sup>1</sup> demonstrate that the purified XPG product specifically nicks a duplex at the border of a single-stranded DNA region. Only the strand with the 5' single-

nucleotide-excision repair reaction? A tentative scheme is depicted in the figure. The proteins XPA and XPE are likely to be involved in damage recognition, possibly with the assistance of transcription factor TFIIH. The latter was originally known as a multisubunit basal transcription factor. But last year it was shown<sup>7</sup> that TFIIH also functions in the context of nucleotide-excision repair, and probably does so in its entirety<sup>8</sup>. Lesion recognition presumably induces a conformational change in the DNA helix. Both DNA 'scissors' require single-stranded DNA, so local unwinding around the lesion before incision would seem to be necessary. The TFIIH complex is an excellent candidate for executing this step, because it displays a bidirectional helicase activity conferred by the XPB and XPD subunits<sup>9</sup>. In humans, the melted region provides a niche for SSB, the complex that binds single-stranded DNA and prevents reannealing. After the incisions by XPG at the 3' end and the RAD1/10 complex at the 5' side, the damage-containing oligonucleotide has to be released. That again may be a task for TFIIH: in basal transcription the complex has been associated with promoter clearance<sup>10</sup>, a mechanistically related activity. Finally, the gap of 27–29 nucleotides is filled in by DNA synthesis and ligation.

This model is an oversimplification. For instance, it does not incorporate the CSA and CSB components, which speed up the repair of the transcribed strand of active genes. Also, no account is taken of the XPC-HHR23B complex and the yeast RAD7 and RAD16 proteins, which are involved in control of the (slower) repair of the remainder of the genome.

The findings reported by O'Donovan *et al.*<sup>1</sup> and Bardwell *et al.*<sup>2</sup> can be placed in a wider clinical perspective. XP-

G is a very rare, heterogeneous form of the disease. Some patients show only features of xeroderma pigmentosum, others additional hallmarks of Cockayne's syndrome. This is a hint that XPG is involved in more than nucleotide-excision repair, and that (depending on the mutation) other processes may become affected. The factors in the transcription/repair complex TFIIH may present a similar case. Mutations in the XPB and XPD subunits can give rise to symptoms of Cockayne's syndrome and trichothiodystrophy. These features may be due to a



stranded end is touched by this endonuclease. It is likely that the yeast equivalent RAD2 does the same<sup>5</sup>. The yeast RAD1/RAD10 duo (and probably its human counterpart, the ERCC1-XPF complex) does essentially the opposite: it nicks the strand of which the 3' end is single<sup>2</sup>. Such a structure is present at the other end of a bubble. These results complete the original observations by Huang *et al.*<sup>6</sup> that the machinery of nucleotide-excision repair makes a dual incision, one five bases 3' and one 21–23 bases 5' of the lesion.

How do these findings fit into the

subtle defect in basal transcription<sup>4,8</sup>. The XPG symptoms suggest a link with the transcription as well<sup>1</sup>, although a firm association with TFIH could not be shown<sup>8</sup>.

The other scissor, RAD1/10, is also engaged elsewhere. It is required in a mitotic recombination pathway<sup>11</sup>, where DNA strands need to be cut as well. ERCC1, the human homologue of RAD10, is one of the few proteins involved in nucleotide-excision repair for which no patients have yet been identified. Obviously, most mutations in this gene may be very severe or give rise to an unexpected clinical picture; indeed, ERCC1-deficient mice show a constellation of abnormalities very different from those characteristic of xeroderma pigmentosum<sup>12</sup>. One might predict that a category of patients will be found that are deficient in nucleotide-excision repair and also have symptoms of a recombination defect. If these striking multiple engage-

ments reflect a general evolutionary strategy of function sharing, then intimate connections between nucleotide-excision repair and cell-cycle control or chromatin dynamics are bound to show up as well. □

Jan H. J. Hoeijmakers and Dirk Bootsma are in the Department of Cell Biology and Genetics, Medical Genetics Centre, Erasmus University, PO Box 1738, 3000 DR Rotterdam, The Netherlands.

1. O'Donovan, A. *et al.* *Nature* **371**, 432–435 (1994).
2. Bardwell, A. J. *et al.* *Science* **265**, 2082–2085 (1994).
3. Cleaver, J. E. *Nature* **218**, 652–656 (1968).
4. Bootsma, D. & Hoeijmakers, J. H. J. *Nature* **363**, 114–115 (1993).
5. Harrington, J. J. & Lieber, M. R. *Genes Dev.* **8**, 1344–1355 (1994).
6. Huang, J.-C., Svoboda, D. L., Reardon, J. T. & Sancar, A. *Proc. natn. Acad. Sci. U.S.A.* **89**, 3664–3668 (1992).
7. Schaeffer, L. *et al.* *Science* **260**, 58–63 (1993).
8. Vermeulen, W. *et al.* *Cold Spring Harb. Symp. quant. Biol.* (in the press).
9. Schaeffer, L. *et al.* *EMBO J.* **13**, 1645–1653 (1994).
10. Goodrich, J. A. & Tjian, R. *Cell* **77**, 145–156 (1994).
11. Schiestl, R. H. & Prakash, S. *Molec. cell. Biol.* **10**, 2485–2491 (1990).
12. McWhir, J. *et al.* *Nature Genet.* **4**, 217–223 (1993).

## EVOLUTIONARY BIOLOGY

# Baby bunting in paternity probe

Kate Lessells

SHOULD a male work as hard to raise his family if there is a chance that he has been cuckolded? The common-sense answer is 'no', but examples of decreasing paternal care with an increasing chance of cuckoldry have proved elusive. The finding by Dixon *et al.* (page 698 of this issue<sup>1</sup>), that male reed buntings feed their brood less when a higher proportion of the chicks are genetically fathered by another male, is therefore one of the first clear demonstrations of a relationship between paternal care and 'paternity' (the proportion of the chicks fathered).

Although male reed buntings defend territories containing the nests of one or more pair-bonded females, they are remarkably unsuccessful in defending their paternity: Dixon *et al.* used single-locus minisatellite fingerprints to show that 55% of the chicks that they looked at were not the offspring of the putative father. Moreover, paternity varies enormously from brood to brood. Some variation is bound to occur: even a true coin is unlikely to produce exactly 50% heads when tossed a few times. But paternity is far more variable than the binomial distribution that would be expected if all reed bunting chicks had the same probability of being illegitimate. In other words, this probability varies from brood to brood.

Whatever the cause of this variation, it provides an ideal opportunity to investi-

gate whether males adjust their brood care to their paternity. Dixon *et al.* looked at the broods of 13 pairs who bred twice in one season. When the change in male feeding rate is plotted against the change in paternity between first and second broods, a clear relationship emerges: males feed heavily cuckolded broods less.

## IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Reed bunting chick feed; the male adult has a black cap.

In contrast, female feeding rate is unrelated to paternity. This confirms that it is the male's lack of genetic relationship with illegitimate chicks that is the important factor, rather than, for instance, illegitimate chicks requiring less food. The paired comparison between first and second broods is also important because it controls for the possibility that particular males are good both at protecting their paternity and providing paternal care, which has been a problem with previous non-experimental studies<sup>2</sup>. Although the

pairwise analysis does not exclude temporary variation in the quality of a male between broods, it does eliminate many confounding variables, such as male age.

One obvious question is what the difference is between reed buntings and other species, for which observational and experimental studies have largely failed to show a relationship between paternal care and paternity<sup>3</sup>. Mathematical models reveal three situations in which paternal care is not expected to vary with paternity.

First, if the relationship between the total fitness of the brood and paternal care is sigmoidal, there is a threshold level of paternity below which the optimal level of paternal care drops abruptly from some to none<sup>4</sup>. But it is only below this threshold that paternal care should be an unvarying nothing; above the threshold paternal care should increase with paternity.

Second, optimal paternal care will not vary much if the relationship between fitness of the brood and paternal care plateaus abruptly<sup>5</sup>: a slight decrease in paternal care will then result in a precipitous drop in brood fitness, while an increase in paternal care will bring no further benefit. However, the exact shape of this relationship is not known for any species, so this offers no explanatory help with the reed buntings.

Lastly, optimal paternal care is not expected to vary with paternity if individual males have consistently low or high paternity<sup>5,6</sup>. A reduction in paternal care by heavily cuckolded males frees time and effort for investment in other reproductive attempts, through re-pairing, making extra-pair copulations or simply surviving to the next breeding season. But if these other attempts are equally cuckolded, paternity will cancel out in an assessment of their value relative to the current brood. Only if the male can expect higher paternity in the future is it worth him reducing care in the current brood. The reed buntings do indeed exhibit high variation in paternity between broods.

Of course, a male can only modify his paternal care in relation to paternity if he has some way of estimating it. How he does this affects how much variation is needed to explain the relationship between paternal care and paternity: if he detects the legitimacy of each chick directly, perhaps through some genetic marker, he can respond to even 'coin-tossing'

1. Dixon, A., Ross, D., O'Malley, S. L. C. & Burke, T. *Nature* **371**, 698–700 (1994).
2. Møller, A. *Anim. Behav.* **36**, 996–1005 (1988).
3. Whittingham, L. A., Dunn, P. O. & Robertson, R. J. *Anim. Behav.* **46**, 139–147 (1993).
4. Whittingham, L. A., Taylor, P. D. & Robertson, R. J. *Am. Nat.* **139**, 1115–1125 (1992).
5. Westneat, D. F. & Sherman, P. *Behav. Ecol.* **4**, 66–77 (1993).
6. Maynard Smith, J. *Anim. Behav.* **25**, 1–9 (1977).
7. Burke, T. *et al.* *Nature* **338**, 249–251 (1989).

Roger Wilmshurst



binomial variation. But reed buntings and other birds<sup>7</sup> seem unable to do this, because some reed bunting males fed broods in which they had no paternity. Instead, they probably use a behavioural cue such as time spent with the female during her fertile period. The male then has no information about binomial variation in paternity, and can only respond to above-binomial variation.

Dixon *et al.* did not carry out the critical analysis of variation in paternity of individual males. But if broods contain four chicks each with a 55% chance of being illegitimate, the probability of a male with

two broods having one with complete and the other with zero paternity is less than 1%. This occurred in three of the 13 cases reported by Dixon *et al.*, so it seems that there is more than enough variation in paternity. It will be interesting to investigate whether those species lacking a relationship between paternal care and paternity also lack above-binomial variation in paternity between broods of individual males. □

Kate Lessells is in the Netherlands Institute of Ecology, Boterhoeksestraat 22, PO Box 40, 6666ZG Heteren, The Netherlands.

## GEOCHEMISTRY

## Fast cleaning in the deep

Louise H. Kellogg

THE Earth's transition zone is something of an enigma. It encompasses the region between two seismic 'discontinuities', the first one 400 km deep in the Earth's mantle, the second 670 km deep, and as the name suggests, it marks the transition between the upper and lower mantle. The transition zone is of interest in part because of the role it plays in the fate of slabs of oceanic crust and lithosphere that are subducted into the mantle at plate margins. Although the path of subducting slabs can be traced to the base of the transition zone by seismology, what actually happens to the slabs within the transition zone and beyond remains a mystery. A current model holds that slab material may pile up for some time in the transition zone until it either descends into the lower mantle in an avalanche<sup>1</sup> or is swept aside by large-scale flow<sup>2</sup>. But do the slabs emerge as mere shadows of their former selves?

In part, the question is how long it takes to wipe out the distinctive chemical signature of slab material. Now Farber *et al.*, on page 693 of this issue<sup>3</sup>, have measured cation diffusion rates experimentally in the high-pressure phases of  $\text{Mg}_2\text{SiO}_4$  thought to be present at transition-zone depths. They conclude that the remarkably high rates of chemical diffusion they find will reduce the time required to mix chemical heterogeneities (such as the subducted oceanic crust) into the transition zone to three-quarters of the time required for mixing heterogeneities in the shallower upper mantle.

Mixing consists of two processes; convective thinning of heterogeneities followed by homogenization by chemical diffusion. According to these new results, chemical diffusion acts on the scale of metres in the transition zone (as opposed to centimetres within the shallow upper mantle). Heterogeneities must be thinned substantially by convection first, but be-

cause the final action of diffusion is more effective in the transition zone, the total time required for mixing is reduced.

The rate of mixing throughout the mantle is undoubtedly controlled primarily by the rate at which convection stirs the material<sup>4</sup>, dispersing heterogeneities into the mantle. Convection stretches and folds heterogeneities, increasing the contact area between the different components being mixed together. The final stage is when diffusion eliminates the finest-scale heterogeneities by smoothing chemical variations.

In the Earth's mantle, convective mixing thins subducted oceanic crust from its initial thickness of 6 km until it is thin enough for diffusion to destroy its distinctive chemical signature. In the shallow upper mantle, diffusive homogenization takes place when the remnants of oceanic crust are only a few centimetres thick. Because diffusion acts only in the last stages of mixing and on such a small length scale, it is usually appropriate to disregard its role in the mixing process.

The new measurements of Farber and colleagues indicate otherwise. In the transition zone, the chemical diffusivity is three orders of magnitude higher and as a result the mixing rate is faster than in the upper mantle. True, the mixing time is only about 25 per cent shorter, even for this enormous change in diffusivity; but that is enough to produce some intriguing predictions.

The higher diffusivity means that the fine-scale structure of the transition zone could be quite different from the structure in the overlying mantle. Very fine-scale heterogeneity (on a scale of 1 cm – 1 m) can exist in the uppermost mantle but only relatively coarse heterogeneity (on scales of 1 m or more) in the transition zone. Heterogeneities less than about a metre thick in the transition zone will be rapidly destroyed by diffusive homogenization.

An interesting additional effect of these results is that, if diffusive equilibrium is more rapidly attained in the transition zone than in the overlying mantle, incompatible components from slabs moving through the transition zone could be deposited more readily into the transition zone than into the overlying mantle. Over the course of hundreds of millions of years, this would add up to a significant enrichment in incompatibles, such as water, in the transition zone. Observations of high concentrations of water in the transition zone have indeed been reported by Nolet and Zielhuis<sup>5</sup>, who observed that seismic shear waves travelled very slowly through a region about 400 km beneath the westernmost Russian Platform. They interpret this and similar anomalies elsewhere as revealing a high water content in the transition zone associated with subduction. The high chemical diffusivity of the spinel and  $\beta$ -phase measured by Farber and colleagues provides a mechanism for homogenization of volatiles from the slab into the transition zone. The increased rates of mixing by chemical diffusion within the transition zone would thus paradoxically lead to development of large-scale heterogeneities between the uppermost mantle and transition zone in the form of layers with different quantities of incompatible elements.

The higher chemical diffusivities could also affect the rheology of transition-zone material. Mantle rock flows by thermally activated creep; increasing the molecular diffusivity by an order of magnitude could decrease the viscosity substantially. If the high diffusivity results in enrichment of incompatibles in the transition zone, the presence of volatiles may also reduce the viscosity. On the other hand, if an excess of garnet were present in the transition zone, that would tend to increase the viscosity<sup>6</sup>. Some recent work<sup>7</sup> does suggest that the transition zone has a lower viscosity than the shallow upper mantle.

The next step will be to reconcile the mineral physics with additional geodynamical, seismological and geochemical data. Further constraints on the volatile content and rheology of the transition zone will help determine the composition and behaviour of this enigmatic and important region of the Earth's mantle. □

Louise H. Kellogg is in the Department of Geology, University of California, Davis, California 95616, USA.

1. Tackley, P. J., Stevenson, D. J., Glatzmaier, G. A. & Schubert, G. *Nature* **361**, 699–704 (1993).
2. Allègre, C. J. & Turcotte, D. L. *Nature* **323**, 123–127 (1986).
3. Farber, D. L., Williams, Q. & Ryerson, F. J. *Nature* **371**, 693–695 (1994).
4. Kellogg, L. H. *Adv. Geophys.* **34**, 1–33 (1993).
5. Nolet, G. & Zielhuis, A. *J. geophys. Res.* **99**, 15813–15820 (1994).
6. Karato, S. *Phys. Earth planet. Inter.* **55**, 234–240 (1989).
7. King, S. D. & Masters, G. *Geophys. Res. Lett.* **19**, 1551–1554 (1992).

# Archaea, Archaea, every where

Gary J. Olsen

In surveying biodiversity, it is all too easy to overlook microorganisms. The reason could not be more obvious — they are very small, and traditional methods for taking a census of them have severe technical limitations. Using methods that count molecules rather than cells, however, DeLong *et al.* (page 695 of this issue<sup>1</sup>) have found that major components of the oceanic picoplankton (in this study, cells smaller than one micrometre) are members of the Archaea — prokaryotic microorganisms that are more closely related to humans than to typical bacteria. In terms of fraction of picoplankton biomass in some cold ocean waters, overlooking the Archaea has been equivalent to surveying one square kilometre of the African savanna and missing over 300 elephants. In terms of biological novelty, overlooking the Archaea would be comparable to omitting all animals from a survey of 'higher organisms'.

The significance of these new findings lies in how they bear on answering several questions. What are the Archaea? How has our view of archaeal abundance been changed? What are the implications of the discovery? And how does it relate to other studies?

Organisms are either eukaryotes (defined by cells with a true nucleus), such as yeasts and humans, or prokaryotes (cells without a nucleus), such as *Escherichia coli* and *Haemophilus influenza*. But molecularly there is an equally important division among prokaryotes: there are two groups — the Archaea and the Bacteria — that are as different from each other (measured by divergence of genes) as are any known organisms<sup>2,3</sup>. Their unique phylogenetic status predicts that the Archaea will have distinctive and diverse physiologies.

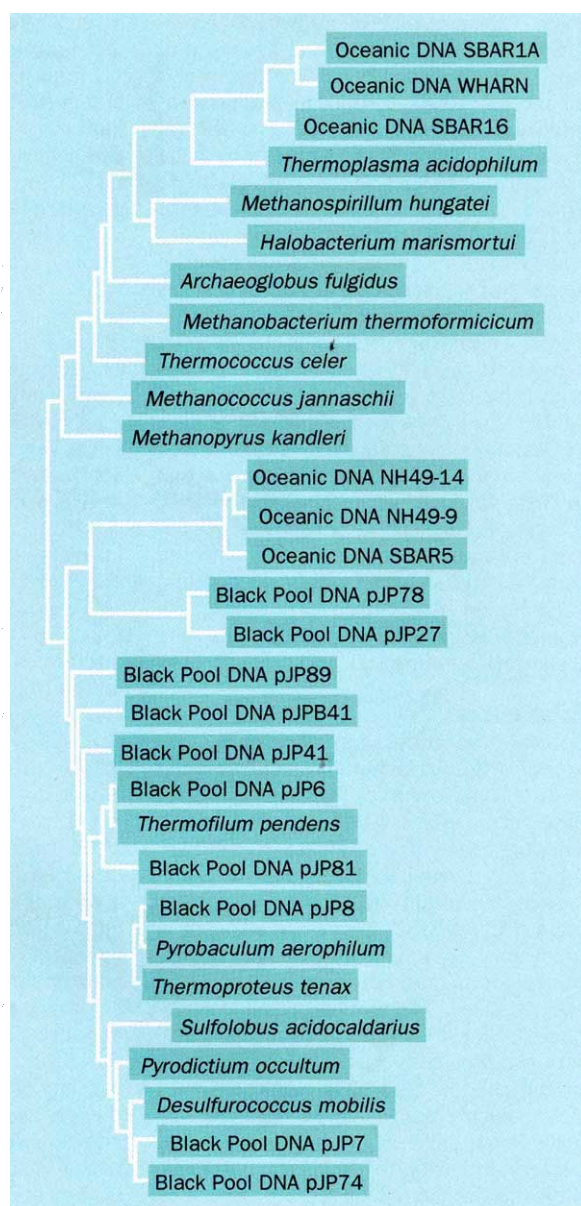
For 15 years after their recognition, the Archaea were generally known only as inhabiting hostile environments (for example acidic hot springs and satu-

rated brines). Because they were not isolated from 'normal' environments (those suitable for humans), Archaea were seen as non-competitive relics, and were not thought to be important in global ecology. However, traditional microbial enumeration methods require cultivation in the laboratory and, as a consequence, the least understood organisms tended to be overlooked. Thus, two years ago, when newer molecular methods were applied, Archaea were unexpectedly found to be widely distributed in the oceans<sup>4,5</sup>. Now DeLong and co-workers have discovered

that the Archaea constitute up to 30 per cent of the submicrometre picoplankton in the cold surface waters of Antarctica and Alaska, and deeper waters elsewhere — no small oversight.

Molecular approaches to microbial ecology provide us with a phylogenetic (genealogical) description<sup>6</sup>, and the ribosomal RNAs of oceanic organisms define two lineages within the Archaea (see figure). What, though, can we say about their physiology? One group is descended from a lineage that spawned all methane-producing organisms, and whose ancestor was almost certainly methanogenic. However, this ancestral lineage has also given rise to extreme halophiles (which have invented photosynthesis without chlorophyll), sulphur reducers, sulphate reducers and some extremely thermophilic heterotrophs. The second archaeal oceanic group belongs to a diverse group called Crenarchaeota, of which all cultured members are thermophiles — a striking contrast with the generally frigid environment of the oceanic Crenarchaeota. In summary, the evolutionary spans separating the oceanic Archaea from their cultured relatives are so great that what the oceanic organisms are doing for a living (that is, their physiology) is, quite frankly, impossible to predict.

Another key aspect of the oceanic Archaea is the degree of diversity within each of the two groups: the RNAs of each group form a cluster of specifically related sequences. If this were the only molecular study of a natural population to observe such 'microheterogeneity'<sup>7</sup>, it might be dismissed as an artefact. However, essentially all studies (including refs 1, 5, 7) that have looked in enough detail have observed some level of variation within the groups present. The differences within a group are substantial, certainly exceeding variation within a species; the divergences are more suggestive of a cluster of species or genera, or even families. It is difficult to extrapolate from observed sequence variations to the associated physiological diversity, making detailed interpretation unsatisfactory. But the image that I prefer is one in which each such grouping has a basic physiological theme, upon which specializations are super-



Phylogenetic tree of Archaea based on 16S ribosomal RNA, showing the approximate phylogenetic positions of sequences from oceanic and Black Pool (Yellowstone National Park) samples. The tree is abstracted from a larger tree available from the Ribosomal Database Project<sup>10</sup>.

- DeLong, E. F., Wu, K. Y., Prézelin, B. B. & Jovine, R. V. M. *Nature* **371**, 695–697 (1994).
- Woese, C. R. & Fox, G. E. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5088–5090 (1977).
- Woese, C. R., Kandler, O. & Wheelis, M. L. *Proc. natn. Acad. Sci. U.S.A.* **87**, 4576–4579 (1990).
- Fuhrman, J. A., McCallum, K. & Davis, A. A. *Nature* **356**, 148–149 (1992).
- DeLong, E. F. *Proc. natn. Acad. Sci. U.S.A.* **89**, 5685–5689 (1992).
- Pace, N. R., Stahl, D. A., Lane, D. J. & Olsen, G. J. *Adv. Microb. Ecol.* **9**, 1–55 (1986).
- Giovannoni, S. J., Britschgi, T. B., Moyer, C. L. & Field, K. G. *Nature* **345**, 61–63 (1990).
- Barns, S. M., Fundyga, R. E., Jeffries, M. W. & Pace, N. R. *Proc. natn. Acad. Sci. U.S.A.* **91**, 1609–1613 (1994).
- Brock, T. D., Brock, K. M., Bely, R. T. & Weiss, R. L. *Arch. Mikrobiol.* **84**, 54–68 (1972).
- Larsen, N. *et al.* *Nucleic Acids Res.* **21**, 3021–3023 (1993).



imposed. Such specializations might include, for example, improved uptake of various carbon sources. Thus, in a complex environment, different species in the group could stably coexist as long as their limiting substrates were distinct.

Although there must be limits to the novelties that will be revealed through the molecular analyses of environments, we are now in the phase of accelerating discovery. In a study of a single hot spring in Yellowstone National Park, Barns and co-workers more than tripled the known phylogenetic diversity among Crenarchaeota (see figure)<sup>8</sup>. This is not due to lack of attempts to cultivate the organisms present by conventional techniques; the first cultured member of the Crenarchaeota was isolated from Yellowstone over 20 years ago<sup>9</sup>, and there have been many studies of the hot springs in the intervening years. So even in extreme environments, where conventional wisdom suggests that population complexity is limited, there is a surprising amount of unrecognized diversity.

The molecular revolution in microbial ecology is providing us with a rich census of microbial life. We can now detect and quantify all significant species in an environment, regardless of our ability to grow them. This is proving to be humbling, in that we are faced with our ignorance of the physiology (and hence the ecology) of many of the organisms that are major components of the environment. Further improving our understanding will require both the improved counting and the continued cultivation and characterization of the organisms we find. □

Gary J. Olsen is in the Department of Microbiology, University of Illinois, Urbana, Illinois 61801, USA.

### Correction

Max Perutz points out that the attribution of the isolation of penicillin to Ernst Chain, which appeared in his obituary of Dorothy Hodgkin (*Nature* **371**, 20; 1994), does not do justice to the other scientists primarily involved. The following details come from a chapter co-written by Hodgkin in *The Chemistry of Penicillin* (Princeton Univ. Press, 1949) and Gwyn MacFarlane's *Howard Florey* (Oxford Univ. Press, 1979).

N. G. Heatley set up mass cultures of *Penicillium notatum* in Howard Florey's laboratory at Oxford, and devised a procedure for extracting penicillin. By January 1941 he had raised production of crude mould filtrate to 500 litres a month, while Chain and E. P. Abraham worked on its purification. The resulting extract, still somewhat impure, was used for the first human trials.

Early in 1942, Abraham and Chain published a brief note in *Nature* (**149**, 328; 1942) reporting the isolation of penicillin in the form of a highly purified barium salt. But it failed to crystallize, and none of the crystalline preparations made later at Oxford proved suitable for Hodgkin's X-ray work to determine its structure. In July 1943, news of the crystallization of the sodium salt of benzyl penicillin at the Squibb Institute for Medical Research, Seattle, reached Oxford; but Hodgkin had to wait until February 1944 before the first 10 milligrams arrived and X-ray studies could commence.

## Inhibitory influences

Martin E. Schwab and Christine E. Bandtlow

PAPERS in the September<sup>1</sup> and October<sup>2</sup> issues of *Neuron* report the surprising observation that myelin-associated glycoprotein (MAG) inhibits neurite growth of selected types of neurons *in vitro*. Myelin-associated glycoprotein is a well-known glial cell-adhesion molecule, which is thought to mediate heterophilic neuronal-glial interactions<sup>3</sup>. The new results, however, suggest that it may be involved in preventing regeneration of lesioned nerve fibres in the brain and spinal cord.

Inhibition and excitation are generally accepted as the basic principle underlying the way the nervous system works. Nevertheless, the idea that axonal growth and guidance are also subject to a fine-tuned balance of growth-promoting and growth-inhibitory signals is quite recent. Several lines of evidence have demonstrated the existence of factors which inhibit neurite growth and lead to growth-cone collapse and repulsion<sup>4,5</sup>. In the context of studies addressing the cell-biological basis for the absence of regeneration of lesioned axons in the adult central nervous system (CNS), a strong inhibitory effect was shown to be associated particularly with the myelinated fibre tract regions of the brain and spinal cord<sup>5</sup>. Indeed, cultured oligodendrocytes or isolated myelin induced growth-cone collapse and growth arrest *in vitro*. *In vivo*, the end of the period in the developing spinal cord when regeneration can occur correlates well with the start of myelin formation; and, in spinal cords made myelin-free experimentally, lesioned axons can regenerate over long distances in young rats and chicks<sup>6,7</sup>.

### Bioassays

Several laboratories have been trying to identify the relevant molecules using neurite outgrowth or growth-cone collapse as bioassays. The task has proved a tough one. The first of the neurite-growth inhibitory activities to be cloned was collapsin, a protein of relative molecular mass ( $M_r$ ) 100K; this is present in a membrane-associated and a secreted form in developing and adult brain, but also in other tissues<sup>8</sup>. Recombinant collapsin induces collapse of growth cones in dorsal root ganglion cells but not in retinal fibres in cell culture. The physiological role of this molecule has yet to be defined. The gene encoding collapsin shows unexpected similarity to the insect guidance molecule fasciclin IV (ref. 9). Antibody experiments have suggested that fasciclin IV functions as an adhesive molecule in grasshopper embryos, although transitory growth-cone collapse has also been observed when sensory neurites come into

contact with epithelial cells expressing the protein<sup>9</sup>.

Attempts to characterize the activities responsible for neurite growth-inhibition in CNS myelin led, in 1988, to semipurified protein preparations of  $M_r$  35K and 250K (refs 10, 11). These constituents are related immunologically but have not yet been fully purified and cloned. McKerracher *et al.*<sup>2</sup> now describe the occurrence of an additional component in spinal cord myelin that inhibits fibre outgrowth of neuroblastoma cells. It turns out to be MAG: the authors found that immunodepletion of MAG from total CNS myelin protein significantly diminished the inhibitory effect on neurite growth.

Mukhopadhyay *et al.*<sup>1</sup> find that recombinant MAG has a strong and selective inhibitory effect on neurite growth of cerebellar and adult dorsal root ganglion (DRG) neurons, whereas neurite outgrowth of DRG from newborns was promoted by the same substrate. In earlier studies, recombinant MAG was shown to bind to newborn mouse DRG cells and to promote neurite outgrowth<sup>12,13</sup>.

The results come as a surprise for several reasons. Myelin-associated glycoprotein has immunoglobulin domains and an Arg-Gly-Asp sequence, and is localized selectively in the periaxonal membrane of Schwann cells and oligodendrocytes, as well as on the outer plasma membrane of Schwann cells; it also occurs as a secreted form in the basement membranes of peripheral nerves<sup>14,15</sup>. It has been viewed as a cell-adhesion molecule that is crucial for Schwann-cell/oligodendrocyte-axon interaction. Antibodies against MAG inhibit oligodendrocyte-oligodendrocyte and oligodendrocyte-neuron adhesion<sup>3</sup>, and transfection of Schwann cells with MAG antisense RNA prevents axonal ensheathment and myelin formation<sup>16</sup>. Given these results, the upshot of the MAG gene-ablation experiments published earlier this year was provocative: myelination of the CNS and the peripheral nervous system (PNS) seemed to be normal, except for subtle structural abnormalities<sup>17,18</sup>.

Myelin-associated glycoprotein occurs in both CNS and PNS, but the two systems differ fundamentally in their regeneration capacity. The rapid clearance of myelin in the lesioned PNS, in contrast to the slow myelin removal in the CNS, could be a partial answer to this discrepancy (although MAG immunoreactivity persists in degenerating/regenerating mouse sciatic nerves<sup>19</sup>). Poor regeneration of sensory axons into anastomosed proximal nerves — with intact myelin — and in mutant mice with delayed myelin degen-

## BACTERIORHODOPSIN

## A nonlinear proton pump

Robert R. Birge

eration has, indeed, been observed<sup>20,21</sup>. The availability of MAG-knockout mice and of good antibodies makes *in vivo* regeneration experiments in the PNS and (especially) the CNS possible. Only these experiments can answer the question of whether MAG is a physiological inhibitor and negative regulator of neurite regeneration.

## Lessons

Two lessons can be learned from the new results. First, a molecule, MAG, which is adhesive and growth-promoting for one type of neuron (perinatal DRG), can inhibit the growth of other neurons or even the same neuron type at a different age (adult DRG). Growth inhibition and growth promotion may, therefore, be exerted by the same molecule, which — probably depending on other ligands, different types of neuronal receptors, or the context in which the receptor signals are interpreted in the growth cone — lead to positive or negative growth responses. Molecules with similar, seemingly divergent effects have, in fact, been characterized already. Members of the tenascin family, for example, can be growth-promoting or neurite growth-inhibitory depending on the neuron type<sup>22</sup>.

Second, *in vitro* experiments can reveal the potential of a given molecule in a well-defined experimental situation. But such experiments have to be complemented by *in vivo* studies to substantiate speculation on their physiological significance. □

Martin E. Schwab and Christine E. Bandtlow are at the Institute for Brain Research, University of Zürich, CH-8029 Zürich, Switzerland.

THE archaeobacterium *Halobacterium halobium* (also called *Halobacterium salinarum*) is one of the oldest organisms on Earth. It populated the salt marshes during the Archaean era more than one billion years ago; it has survived through ten evolutionary eras and seven mass extinctions; and today it is evident to air

used by *H. halobium* to synthesize ATP from inorganic phosphate and ADP. The discovery prompted extensive study of bacteriorhodopsin, and much has been learned during the past decade and a half (for a review see ref. 3).

What Tributsch and Bogomolni have now seen, though, is rather unexpected.

They prepared envelope vesicles of the archaeobacterium and observed sustained photo-oscillations of a fluorescent pH probe under certain illumination regimes, pH values and protein concentrations. They identified bacteriorhodopsin as the source of the oscillations by preparing liposomes containing only the protein and the probe (Fig. 1) and observing oscillations in both fluorescence and pH, examples of which are shown in Fig. 2. These, suggest the authors, are evidence that bacteriorhodopsin incorporates a synchronized, positive feedback mechanism in regulating its proton pumping behaviour<sup>1</sup>.

Let us concentrate on the liposome experiments, because these are less complicated. The liposomes were constructed by mixing purified azolectin lipids in saline solution with purified bacteriorhodopsin. Via a mechanism that is well characterized but not fully understood, reconstituted liposomes containing bacteriorhodopsin orient the protein so that the proton

pumping is one-way, as shown in Fig. 1. The liposomes were irradiated by a filtered tungsten lamp to provide incoherent excitation, at an intensity ranging from 0.22 to 220 mW cm<sup>-2</sup>. Both the relative fluorescence from the cyanine dye and the pH of the solution oscillated (Fig. 2) — but intriguingly, the fluorescence oscillated at twice the frequency of the pH.

Tributsch and Bogomolni suggest that the cyanine dye emits fluorescence in response to field strength, irrespective of the sign of the field vector, so a proton going in either direction through the membrane will increase fluorescence. In contrast, the micro-pH electrode responds to a change in hydrogen ion concentration.

There are two possible explanations for

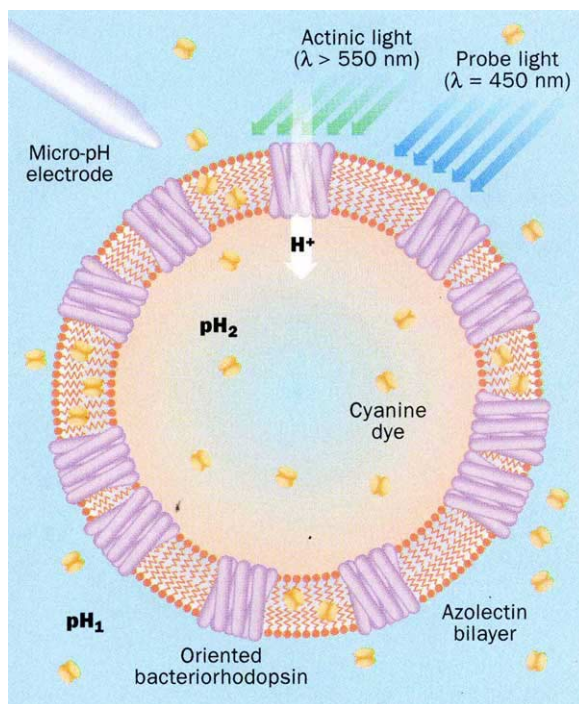


FIG. 1 Bacteriorhodopsin is a membrane-bound protein that has seven  $\alpha$ -helical segments that span the membrane. The liposome experiments carried out by Tributsch and Bogomolni<sup>1</sup> are schematically shown (see text). The protein was photoactivated with wavelengths above 550 nm whereas the cyanine dye probes were excited at 450 nm and the fluorescence was measured at 506 nm. The dye partitions into the membrane, and it is the membrane fraction which responds to the field gradient produced by proton translocation through the membrane.

travellers entering San Francisco, being responsible for the deep purple colour of the salt flats. Nevertheless, *H. halobium* and its 'purple membrane' continue to surprise us, as a new study by Tributsch and Bogomolni<sup>1</sup> illustrates.

When the concentration of oxygen drops below the level needed to sustain oxidative phosphorylation, the organism grows a purple membrane which contains the protein bacteriorhodopsin. In the early 1970s, Oesterhelt and Stoekenius<sup>2</sup> discovered that on absorbing light, the protein undergoes a photocycle which pumps protons across the membrane, with a net transport from the inside (cytoplasmic) to the outside (extracellular) of the membrane. The resulting pH gradient generates a proton-motive force which is

- Mukhopadhyay, G., Doherty, P., Walsh, F. S., Crocker, P. R. & Filbin, M. Y. *Neuron* **13**, 757–767 (1994).
- McKerracher, L. et al. *Neuron* **13**, 805–811 (1994).
- Poltorak, M. et al. *J. Cell Biol.* **105**, 1893–1899 (1987).
- Fawcett, J. W. *Trends Neurosci.* **16**, 165–167 (1993).
- Schwab, M. E., Kapfhammer, J. P. & Bandtlow, C. E. *A. Rev. Neurosci.* **16**, 565–595 (1993).
- Savio, T. & Schwab, M. E. *Proc. natn. Acad. Sci. U.S.A.* **87**, 4130–4133 (1990).
- Keirstead, H. S., Hasan, S. J., Muir, G. D. & Steeves, J. D. *Proc. natn. Acad. Sci. U.S.A.* **89**, 11664–11668 (1992).
- Luo, Y., Raible, D. & Raper, J. A. *Cell* **75**, 217–227 (1993).
- Kolodkin, A. L. et al. *Neuron* **9**, 831–845 (1992).
- Caroni, P. & Schwab, M. E. *J. Cell Biol.* **106**, 1281–1288 (1988).
- Caroni, P. & Schwab, M. E. *Neuron* **1**, 85–96 (1988).
- Johnson, P. W. et al. *Neuron* **3**, 377–385 (1989).
- Schneider-Schaulies, J., Kirchhoff, F., Archelos, J. & Schachner, M. *Neuron* **7**, 995–1005 (1991).
- Trapp, B. D., Andrews, S. B., Cootauco, C. & Quarles, R. *J. Cell Biol.* **109**, 2417–2426 (1989).
- Martini, R. & Schachner, M. *J. Cell Biol.* **103**, 2439–2488 (1986).
- Owens, G. C. & Bunge, R. P. *Neuron* **7**, 565–575 (1991).
- Li, C. et al. *Nature* **369**, 747–750 (1994).
- Montag, D. et al. *Neuron* **13**, 229–245 (1994).
- Martini, R. & Schachner, M. *J. Cell Biol.* **106**, 1735–1746 (1988).
- Brown, M. C., Lunn, E. R. & Perry, V. H. *Eur. J. Neurosci.* **3**, 1366–1369 (1991).
- Brown, M. C., Perry, V. H., Hunt, P. S. & Lapper, S. R. *Eur. J. Neurosci.* **6**, 420–428 (1994).
- Schachner, M., Taylor, T., Bartsch, U. & Pesheva, P. *Perspect. dev. Neurobiol.* **2**, 33–41 (1994).



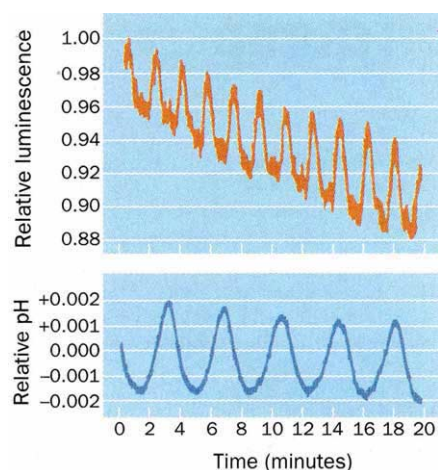


FIG. 2 When bacteriorhodopsin is incorporated into liposomes in the presence of cyanine fluorescent dyes, continuous radiation produces oscillation in the fluorescence from the dye (red trace) and in the relative pH (blue trace) under certain conditions. Data reproduced with permission from ref. 1 (bacteriorhodopsin concentration  $2.7 \mu\text{g ml}^{-1}$ , actinic ( $\lambda > 550 \text{ nm}$ ) intensity  $55 \text{ mW cm}^{-2}$ ,  $\text{pH}_{\text{dark}} = 6.8$ ).

the observed oscillations. The first, and trivial, explanation is that the permeability of the membrane is undergoing repeated breakdown and repair, but this is unlikely given the two decades in field gradient under which oscillations were observed. The more exciting possibility is that the photoinduced proton pumping of bacteriorhodopsin includes a cooperative process involving synchronous, positive feedback.

Engineers have long recognized that frequency encoding is preferable to amplitude encoding with respect to resistance to noise and precision of control — for instance, frequency modulation (FM) is preferable to amplitude modulation (AM) for radio reception and bandwidth. Raap and co-workers have proposed that the oscillatory behaviour that characterizes many biochemical control networks is the biological analogue of the above engineering principle<sup>4</sup>. For example, some insects emit pheromones with a periodic frequency that is species-dependent. Not only does the oscillatory emission improve sensitivity for the receiver, but only members of the same species are tuned to respond to the frequency, a process known as 'odour song'.

Ross and Schell, though, suggest instead that oscillatory behaviour in biological systems is selected through evolution because it enhances thermodynamic efficiency<sup>5</sup>. The idea is that the rate of a chemical reaction is inversely proportional to the free energy barrier ( $\Delta G$ ) associated with the reaction step. If the barrier increases as a function of the reaction, which is often the case in biological reactions, then oscillatory behaviour can enhance efficiency by adjusting the phase

between the fluxes and the forces. The dissipation is reduced if the rate is a maximum at minimum  $\Delta G$  and a minimum at maximum  $\Delta G$ . Nonlinearities (positive feedback) in the reaction mechanism are essential for maintaining the proper phase relationships. The engineering analogue is to enhance the efficiency of an electronic device by minimizing the phase difference between voltage and current.

Regardless of origin, oscillations in biological systems are ubiquitous, and over 140 examples have been examined<sup>4,5</sup>. All such systems can be interpreted as satisfying one of the above two models, and many appear to satisfy both. So there seems to be an evolutionary force that favours oscillatory behaviour when the available energy approaches the barrier to the process. In cases like these, nonlinear behaviour is selected to provide greater efficiency or resistance to noise.

This is essential in the archaeobacterium, as the primary photochemical event of bacteriorhodopsin stores only about 12 kcal per mol (about 50 kJ per mol), which leaves less than 4 kcal per mol of excess energy under nominal conditions<sup>6</sup>. The observed oscillations may represent the use of autocatalytic behaviour to optimize the efficiency. Perhaps, for instance, the protein adjusts its reverse proton permeability in response to the bulk properties of the solution so that the phase relationship between the proton pumping process and the field gradient exhibits minimal dissipation (in this case, when they are out of phase). This process is analogous to the use of a capacitor on an electric motor so that the voltage and the current are maintained in phase for greatest efficiency.

The mechanistic origins of bacteriorhodopsin's oscillatory behaviour remain to be discovered, and this unexpected autocatalytic behaviour should tell us something about the proton pumping mechanism itself. If the proton pumping process really is being optimized by a synchronous nonlinear mechanism, it only goes to show that the key optimization principles of engineering have already been implemented in biological systems. □

Robert R. Birge is in the Department of Chemistry and the W. M. Keck Center for Molecular Electronics, Syracuse University, Syracuse, New York 13244, USA.

1. Tributsch, H. & Bogomolni, R. A. *Chem. Phys. Lett.* **227**, 74–78 (1994).
2. Oesterholt, D. & Stoekenius, W. *Nature New Biol.* **233**, 149–152 (1971).
3. Birge, R. R. *Biochim. biophys. Acta* **1016**, 293–327 (1990).
4. Raap, P. E., Mees, A. I. & Sparrow, C. T. *J. theor. Biol.* **90**, 531–544 (1981).
5. Ross, J. & Schell, M. A. *Rev. Biophys. biophys. Chem.* **16**, 401–422 (1987).
6. Birge, R. R. *et al. J. Am. chem. Soc.* **113**, 4327–4328 (1991).

DAEDALUS

## Clearing the air

HUMAN bodily odours are widely feared. The soap and cosmetics industries sell many of their products not as removers of dirt, but as destroyers of smell. Daedalus points out that the volatile smelly emissions from the human frame must be very small — perhaps a few milligrams a day. This tiny outflow could easily be mopped up by the sort of absorptive charcoal found in smell-absorbing kitchen ventilators. A carbon-fibre cloth is already widely used in gas masks and smoke filters; in the form of smart black underwear it should absorb body effluvia perfectly. A combined weave of carbon and polyester might be stronger and more practical.

But even the best absorption filter soon gets saturated. Carbon-based anti-smell underclothes could, of course, be regenerated by heating them in a vacuum; but Daedalus has a better idea. He recalls the chemical catalysts widely used to oxidize fumes and smells in industrial exhaust air. They are porous materials such as charcoal or ceramic, loaded with some finely divided catalytic metal. They usually run hotter than  $100^\circ\text{C}$ , but some will work below this. DREADCO's chemists hope to perfect catalysts which will work slowly even at body heat. In an absorptive garment, they would slowly burn up entrapped odour molecules, regenerating the absorbent automatically.

Daedalus is even testing a cloth knitted from polyester and ultrafine platinum wire. This vastly expensive fabric is made for oxidizing ammonia, but it should also burn up the related amines of human effluvia. As lingerie it would provide the ultimate in costly, hygienic elegance and chic.

DREADCO's 'Catawear' should fill a long-smelt want. The fear of being unwittingly offensive to our fellows, played on so subtly by the advertisers, will be vanquished at last. Personal cleanliness will cease to obsess us, and will be downgraded to the occasional removal of annoying adherent matter. Catawear will subvert the sales, not just of deodorants, but even of soap itself.

More revolutionary still, catalytic clothes may never need washing. Volatile organic stains would be burned up, and even solid grime should slowly disappear. Each grime particle would be slowly oxidized at its points of attachment to the catalyst, whereupon it would simply fall off. Hanging in the wardrobe or bundled in the drawer, Catawear garments will slowly clean themselves. They may even scavenge grime and dead cells from the wearer's own skin, reducing sales of soap even further.

David Jones

# Crystals of Hg superconductors

SIR — Solid-state high-pressure techniques are very useful for synthesizing polycrystalline copper oxide superconductors, including the recently discovered  $\text{HgBa}_2\text{Ca}_{n-1}\text{Cu}_n\text{O}_{2n+2+\delta}$  ( $\text{Hg-12}(n-1)n$ ) and the infinite-layer compound  $\text{Ca}_{1-x}\text{A}_x\text{CuO}_2$  (where A is an alkaline earth metal)<sup>1,2</sup>. High pressure, produced

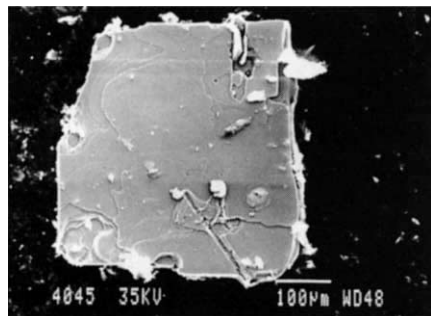


FIG 1.  $\text{HgBa}_2\text{Ca}_3\text{Cu}_4\text{O}_x$  single crystal with characteristic growth steps on the surface (P. Wägli).

by a solid medium such as pyrophyllite, prevents mercury evaporation in the former and stabilizes the structure in the latter. But the sample mass is limited up to several hundred milligrams, there is no free space to grow single crystals, the pressure and temperature distribution are not homogeneous and the partial oxygen pressure cannot be controlled. We have developed a gas-phase high-pressure technique which circumvents these problems and which has allowed us to grow single crystals of  $\text{Hg-12}(n-1)n$  of millimetre size.

The main advantages of this method compared with the solid-state pressure medium technique are: (1) the use of an Ar gas atmosphere with a defined partial oxygen pressure leaves free space for the single crystal growth; (2) pressurized Ar gas acts as an encapsulation of the sample, preventing evaporation of Hg; (3) the temperature gradient can be controlled in a three-zone furnace; and (4) the sample volume can be as large as several  $\text{cm}^3$ .

The experimental high pressure system consists of a "Unipress" gas compressor 15-kbar and a piston-cylinder pressure chamber with an internal three zone Kanthal furnace<sup>3,4</sup>.

The  $\text{Hg-12}(n-1)n$  compounds melt peritectically, but at ambient pressure decompose before melting and the volatile components evaporate. Because of the high density of Ar gas at this pressure the evaporation of Hg is strongly suppressed. Application of a flux with a lower melting point allows growth of single crystals below the peritectic decomposition temperature. Mixtures of the stoichiometric  $\text{Hg-1223}$  with 10 atom %  $\text{PbO}$  or 50 atom %  $\text{BaCuO}_2$  plus  $\text{CuO}$  have been

used as a flux in several crystal-growth experiments. Parameters of the crystal growth process were:  $p_{\text{Ar}} = 10$  kbar,  $T_{\text{max}} = 1,040 - 1,070$  °C, a cooling rate of 5–20 °C per hour down to 1,020 °C followed by a cooling rate of 10 °C per min to 20 °C.

As a result, single crystals of the  $\text{Hg-12}(n-1)n$  compounds (with  $n = 3, 4$  and 5) have been obtained. For the  $\text{Hg-1234}$  single crystals a  $T_c$  of 129 K was observed, which is higher than 126 K reported previously (J. Capponi, unpublished data). In Fig. 1, the  $\text{Hg-1234}$  single crystal is shown with characteristic growth steps on the surface. A separation of the crystals from the flux is very difficult, so we have obtained pieces only of size up to  $0.5 \times 0.5 \text{ mm}^2$ . We expect to optimize the crystal growth process to obtain larger crystals.

A  $\text{Hg-1234}$  crystal was measured on an X-ray four-circle single-crystal diffractometer, using  $\text{MoK}\alpha$  radiation ( $\lambda = 0.7103$  Å). The space group is  $P4/mmm$  and the lattice parameters are  $a = b = 3.8478$  Å,  $c = 18.982$  Å (Fig. 2). The assumed four-layer structure was confirmed, the refinements converged at  $R = 0.064$ . In the  $c$ -direction, the structure contains four  $\text{CuO}_2$  layers separated by Ca atoms. They are sandwiched by two Ba–O layers which include one Hg–Pb–O plane.

The application of the gas high pressure technique to  $\text{Hg-12}(n-1)n$  family will allow optimization of the synthesis and crystal growth in thermodynamically well defined conditions.

During the preparation of this manuscript, a paper concerning structural investigations of  $\text{Hg-1212}$  and  $\text{Hg-1223}$  single crystals of a maximum size 30  $\mu\text{m}$  obtained by recrystallization of a powder has been published<sup>5</sup>. More details concerning our high pressure growth method and the structure and properties of  $\text{Hg-12}(n-1)n$  and  $\text{CaCuO}_2$  infinite layer

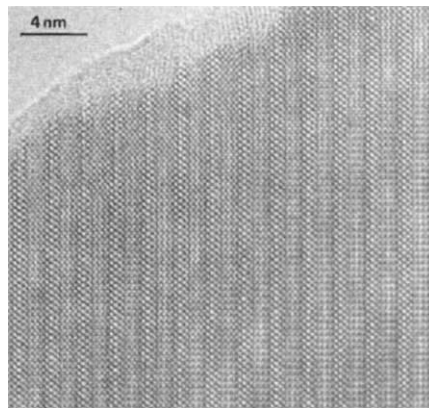


FIG 2. High-resolution transmission electron micrograph of a perfect structure of the  $\text{Hg-1234}$  single crystal (R. Wessiken and H.-U. Nissen). Scale bar, 4nm.

superconducting single crystals will be presented elsewhere<sup>6</sup>.

J. Karpinski

H. Schwer

I. Mangelschots

K. Conder

Laboratorium für Festkörperphysik,  
ETH 8093-Zürich, Switzerland

A. Morawski

T. Lada

A. Paszewin

High Pressure Research Centre

"Unipress", Polish Academy of Science,  
01-142-Warsaw, Poland

1. Antipov, E. et al. *Physica C* **215**, 1–10 (1993).
2. Azuma, M., Hiroi, Z., Takano, M., Bando, Y. & Takeda, Y. *Nature* **356**, 775–776 (1992).
3. Morawski, A., Lada, T., Pachla, W., Paszewin, A. & Molinski, R. A., *Proc. ICMAS Conf. 1991* (ed. Niku-Lari, A.) 361–365, (IIT International, Paris, 1991).
4. Karpinski, J. Jr. & Porowski, S. *J. Cryst. Growth*, **66**, 1–10 (1984).
5. Finger, L. et al. *Physica C* **226**, 216–221 (1994).
6. Karpinski J. et al. *Physica C* (in the press).

## Plant survival

SIR — Grabherr *et al.* in their Scientific Correspondence<sup>1</sup> present a grim prognosis for montane plants trying to keep up with global warming. But the situation could be even worse. Climate change will probably involve both temperature and precipitation changes at most locations<sup>2</sup>. When two physical factors change, the range of a species may not only shift, but may also change in size because the upper and lower limits of a species are not determined by the same factor<sup>3</sup>. For example, the range of intertidal barnacles is limited by desiccation from above and inter-specific competition from below<sup>4</sup>. For plants along gradients of latitude or elevation, temperature often determines the upper limit<sup>5–7</sup>, whereas moisture<sup>5</sup> or carbon balance<sup>8</sup> (at least partly dependent on moisture availability) often determine the lower limit. Wherever temperature increases and moisture decreases (a frequent occurrence in many models<sup>2</sup>), the range of a species is likely to shrink as well as shift. Therefore, predictions about the effect of climate change on plant distributions need to take both temperature and moisture into account.

Owen D. V. Sholes

Department of Natural Sciences,

Assumption College,

Worcester, Massachusetts 01609, USA

1. Grabherr, G., Gottfried, M. & Pauli, H. *Nature* **369**, 448 (1994).
2. Peters, R. L. & Lovejoy, T. E. (eds) *Global Warming and Biological Diversity* (Yale Univ. Press, New Haven, 1992).
3. Kareiva, P. M., Kingsolver, J. G. & Huey, R. B. (eds) *Biotic Interactions and Global Change* (Sinauer, Sunderland, Massachusetts, 1993).
4. Connell, J. H. *Ecology* **42**, 710–723 (1961).
5. Pearson, G. A. *Ecology* **9**, 145–160 (1930).
6. Davis, M. B. & Zabinski, C. In *Global Warming and Biological Diversity* (eds Peters, R. L. & Lovejoy, T. E.) 297–308 (Yale Univ. Press, New Haven, 1992).
7. Pigott, C.D. & Huntley, J. P. *New Phytol.* **87**, 817–839 (1981).
8. Bunce, J. A., Chabot, B. F. & Miller, L. N. *Bot. Gaz.* **140**, 288–294 (1979).



## Source of unique tumour antigens

SIR — there is mounting enthusiasm for the idea that antigens in malignant cells might be used as vaccines to induce tumour-specific cell-mediated immunity (see ref. 1 for review). A fine example of the power of T cells to destroy solid masses of malignant tissue in man has recently been demonstrated by Papadopoulos et al.<sup>2</sup>, who transferred mature donor T cells to immunosuppressed recipients of bone marrow grafts in whom Epstein-Barr virus-associated lymphoproliferative disease had developed.

Work has focused on various sources of antigens, including viral<sup>2,3</sup>, fetal and tissue-specific antigens<sup>1</sup>, and point mutations<sup>4</sup>. Here we suggest an additional, potentially powerful source of tumour antigens that may arise by frameshift mutations. Frameshift mutations give rise to new segments of unique protein. Although these are smaller than most viral proteins, they are more likely than point mutations to contain sequences that will bind to HLA molecules and be presented at the cell surface to cytotoxic T lymphocytes<sup>5</sup>.

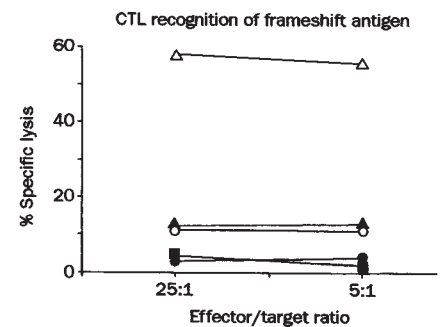
Frameshift mutations are particularly common in carcinoma of the colon<sup>6,7</sup>, which is one of the commonest malignancies in the western world (there are about 160,000 new cases each year in the United States<sup>8</sup>). They also occur in gastric and pancreatic carcinoma<sup>9,10</sup>, and occasionally in other malignancies<sup>11</sup>. In carcinomas of the colon the adenomatous polyposis coli gene (APC) is mutated early in most cases, being detectable in pre-malignant colonic polyps. Nearly all the mutations in APC give rise to a truncation of the encoded protein<sup>6,7</sup>. Approximately half of the mutations are single base changes that result in the formation of a 'stop' codon, and these are not of immunological interest. The other half arise by short deletions

or insertions that alter the reading frame<sup>7</sup>. In these cases, beyond the mutation a potentially unique protein sequence will be read off the second or third reading frames until a stop codon is encountered.

We have analysed the series of 43 cases of carcinoma of the colon described in ref. 7. Of these, 21 had frameshift mutations in the APC gene that we calculate would give rise to new sequences of up to 56 amino acids, with a mean length of 16.1 residues. The 11 cases with predicted new sequences of 16 residues or longer are listed in the table. Several of these new reading frames contain regions of sequence that are likely to bind common HLA class I molecules<sup>12</sup>, and we have confirmed this for the underlined sequences by demonstrating their ability to induce the assembly of HLA A2 class I molecules *in vitro*<sup>13</sup> (data not shown).

As the mutations tend to cluster<sup>7</sup>, certain new reading frames are represented more frequently than others. To see whether these sequences could be immunogenic, we cloned the most commonly expressed new reading frame, expressed it in recombinant vaccinia, and vaccinated BALB/c mice. These animals produced a vigorous cytotoxic T-lymphocyte (CTL) response to the sequence in bold type, which was antigen-specific and MHC (K<sup>d</sup>) restricted (see figure). This result demonstrates the viability of using these sequences as immunogens.

Much work is required to establish the value of these sequences as tumour vaccines. Although the 53-amino-acid fragment containing this epitope was apparently correctly processed in L cells (see figure), and the colonic tumour-derived cell line LoVo (data not shown) after infection with recombinant vaccinia, it is not known whether processing of the full-length protein would occur in malignant colonic epithelial cells *in vivo*. In addition, such cells may be highly mutable<sup>14</sup>, and may lose the function of genes required for antigen presentation<sup>15,16</sup>. In summary, frameshift muta-



Recognition of a frameshift antigen by K<sup>d</sup>-restricted CTL. BALB/c mice were primed with 10<sup>7</sup> PFU recombinant vaccinia virus (M-53-Q-VAC) encoding a 53-amino-acid sequence (MAPVIFQIALDKPCHQAEVKHLHLLKQLKPSEKYLKIKHLLKRERVDLSKLQ) from the second reading frame of human APC containing the sequence KYLKIKHLL (in bold type in the table). Spleen cells were restimulated and maintained *in vitro* with the peptide KYLKIKHLL derived from this reading frame, and tested in a standard <sup>51</sup>Cr release assay at the effector target ratios shown on L/Db target cells<sup>5</sup>. Target cells were either uninfected (solid squares); infected with M-53-Q-VAC (solid circles); infected with K<sup>d</sup>-VAC alone (solid triangles); infected with M-53-Q-VAC and K<sup>d</sup>-VAC (to provide the class I restriction element) (empty triangles); infected with K<sup>d</sup>-VAC and Ub-RNP-VAC (encoding a short lived influenza nucleoprotein as a control) (empty circles). K<sup>d</sup>-VAC was kindly provided by J. Yewdell.

tions offer an additional source of tumour-specific antigens that are potentially more immunogenic than point mutations, but less so than viral antigens. They are relatively easy to detect, are common in carcinoma of the colon and other gastrointestinal malignancies, and should give rise to protein fragments that are likely to bind to the common HLA class I molecules in a significant proportion of cases.

Alain Townsend, Claes Öhlén  
Mike Rogers, Jon Edwards  
Siddhartha Mukherjee, Judy Bastin  
Institute of Molecular Medicine,  
University of Oxford,  
John Radcliffe Hospital,  
Oxford, OX3 9DU,  
UK

NEW SEQUENCES GENERATED BY FRAMESHIFT MUTATIONS IN THE APC GENE

| Codon | Mutation | New Sequence                                                       | Size |
|-------|----------|--------------------------------------------------------------------|------|
| 298   | 2bp del  | SSST/LCTSKADKSSGNQGGNGVFIVNAWYS                                    | 27aa |
| 540   | 1bp del  | SED/LTAGYCKCFEEFVLASRCK                                            | 18aa |
| 1068  | 4bp del  | EQRQ/GIKVQLILFILRALMINTSSSNHIL<br>DSRNVFLHTGHGPEMVQKQIEWVLIMELIKM  | 56aa |
| 1353  | 8bp del  | HKAV/FRSEISLQKWCSDTQKST                                            | 18aa |
| 1398  | 1bp del  | DSFE/SVRLPAPFRVNAHAVEV                                             | 16aa |
| 1420  | 1bp del  | IISP/VIFQIALDKPCHQAEVKHLHLLK<br><u>OLKPSEKYLKIKHLL</u> KRERVDLSKLQ | 51aa |
| 1439  | 1bp del  | RSKT/LHLLKQLKPSEKYLKIKHLL<br>LKRERVDLSKLQ                          | 33aa |
| 1446  | 10bp del | PPQT/GEKYLKIKHLLKRERVDLSKLQ                                        | 23aa |
| 1488  | 1bp del  | DADT/YILPRKVLQMDFLVHPA                                             | 18aa |
| 1490  | 1bp del  | DTLL/LPRKVLQMDFLVHPA                                               | 16aa |
| 1493  | 11bp del | LHFA/SRWIFLFIQPECSEPR                                              | 16aa |

The data are derived from ref. 7. The sequences underlined were found to stabilize the assembly of HLA A2 as described<sup>13</sup>, the sequence in bold induced K<sup>d</sup> restricted CTL in BALB/c mice.

- Pardoll, D.W. *Nature* **369**, 357–358 (1994).
- Papadopoulos, E.B. et al. *N. Engl. J. Med.* **330**, 1185–1191 (1994).
- Feltkamp, M.C.W., et al. *Eur. J. Immun.* **23**, 2242–2249 (1993).
- Peace, D.J. et al. *J. exp. Med.* **179**, 473–479 (1994).
- Townsend, A.R.M., Gotch, F.M. & Davey, J. *Cell* **42**, 457–467 (1985).
- Nishisho, I. et al. *Science* **253**, 665–669 (1991).
- Miyoshi, Y. et al. *Hum. molec. Genet.* **1**, 229–233 (1992).
- Winawer, S.J. *N. Engl. J. Med.* **328**, 1416–1417 (1993).
- Hori, A. et al. *Cancer Res.* **62**, 6696–6698 (1992).
- Hori, A. et al. *Cancer Res.* **52**, 3231–3233 (1992).
- Harris, C.C. & Hollstein, M. *N. Engl. J. Med.* **329**, 1318–1327 (1993).
- Falk, K. et al. *Nature* **351**, 290–294 (1991).
- Townsend, A. et al. *Cell* **62**, 285–295 (1990).
- Aaltonen, L.A. et al. *Science* **260**, 812–816 (1993).
- Smith, M.E.F., Marsh, S.G.E., Bodmer, J.G., Gelsthorpe, K. & Bodmer, W.F. *Proc. natn. Acad. Sci. U.S.A.* **86**, 5557–5561 (1989).
- Kaklamani, L. et al. *Am. J. Path.* (In the press).

## Rapid growth at deep-sea vents

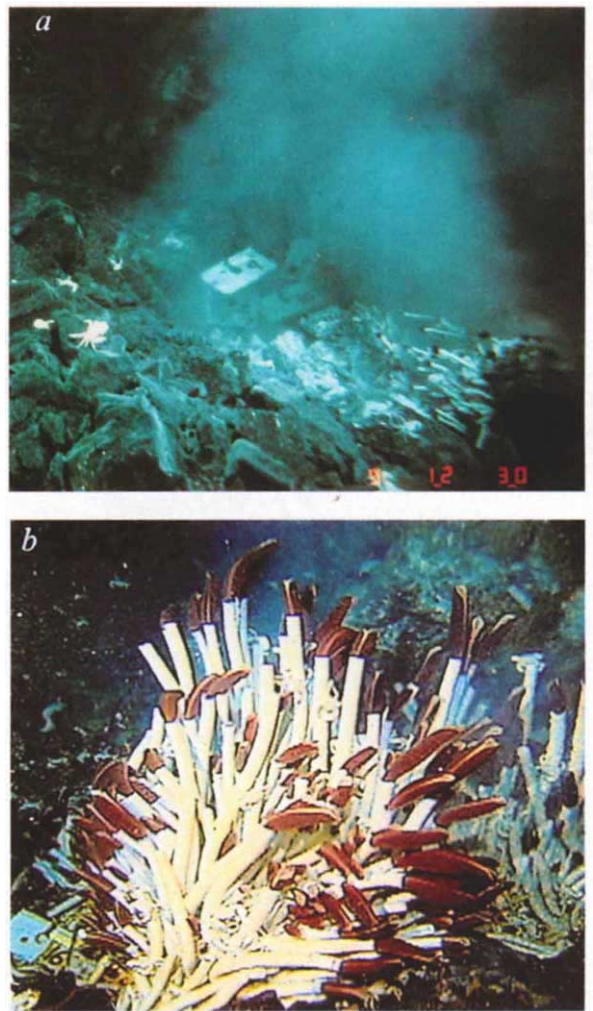
**SIR** — Rates of biological processes at deep-sea hydrothermal vents have been the subject of considerable research since the initial discovery of vent ecosystems in 1977. Previous studies have suggested that the environmental unpredictability, transient nature and high biological production of hydrothermal systems favour continuing reproduction, rapid recruitment, accelerated growth and a tolerance to environmental perturbations<sup>1-3</sup>. Although there is some evidence suggesting that certain vent organisms possess many of these attributes<sup>1-5</sup>, numerous fundamental questions remain concerning growth, mortality, reproduction, larval dynamics, recruitment and many other physiological characteristics of the vast majority of organisms inhabiting vent environments. Our knowledge of rates of colonization and growth of organisms endemic to deep-sea, hydrothermal habitats is particularly limited and based largely on results obtained from relatively few direct and indirect studies in which various assumptions have been made<sup>2,3,5</sup>. Here we present unambiguous direct evidence of extremely rapid rates of colonization and growth of vestimentiferan tube worms, one of the dominant and most conspicuous members of many deep-sea hydrothermal vent communities.

In April 1991, the deep submergence vehicle *Alvin* visited an area of recent volcanic activity between 9°45'N and 9°52'N along the crest of the East Pacific Rise at a depth of approximately 2,500 m. At that time the axial summit caldera of the ridge crest was characterized by extensive new lava flows; numerous high-temperature (>360 °C) vents (with superheated fluids exiting directly from fissures in bare basalt); large quantities of low-temperature (<30 °C), diffuse flow; 'snowstorms' of flocculent, biogenic debris in the water column; and extensive, thick (up to about 5 cm), white bacterial mats covering the basalt sea floor<sup>6</sup>. No characteristic vent megafauna were observed or photographically documented in those areas of the axial summit caldera that had become hydrothermally active as a result of the

1991 volcanic eruptive event(s).

*Alvin* returned to the area in March 1992 to find a marked reduction in the quantity of bacteria throughout the region; low-temperature, hydrothermal venting and associated bacterial mats were restricted to localized, discrete sites within the axial summit caldera. Numerous venting fissures, previously uninhabited by vent megafauna, were colonized by extensive populations of the vestimentiferan tube worm *Tevnia jerichonana* with tubes measuring up to 30 cm long (*a* in the figure). Collections made at several of these vent areas yielded no specimens of the giant tube worm, *Riftia pachyptila*, and no specimens of this vestimentiferan were detected in any of the numerous video images taken at these newly formed vent sites. Numbered polyethylene markers were deployed on the sea floor during the March 1992 expedition in several of these vent areas to facilitate recognition of specific locations; markers 9, 10 and 16 (see figure *a*) were placed directly beside one another in a population of *T. jerichonana* vestimentiferans inhabiting an active, low-temperature venting fissure within the axial summit caldera at 9°50.35'N.

In December 1993, *Alvin* returned to the same localities visited in March 1992 and found dramatic changes in biological community structure at the various nascent hydrothermal vent sites. In most of those areas that were still hydrothermally active, large numbers of giant tube worms (*R. pachyptila*) had settled and grown, with tubes of many of the individual organisms having reached lengths in excess of 1.5 m. Figure *b* depicts a large colony of *R. pachyptila* enshrouding both the *T. jerichonana* population and markers 9 and 10. Colonies with a similar abundance and size range of these large vestimentiferans were present at numerous other nearby vent sites that were uninhabited by this species in March 1992; individual organisms within these colonies were



A fissure from which hydrothermal fluid is venting at a depth of 2,504 m along the crest of the East Pacific Rise (Lat. 9° 50.3'N; long. 104° 17.4'W). No characteristic vent megafauna were present at this fissure when it became hydrothermally active during a volcanic eruptive event in April 1991. By March 1992 (*a*), the fissure was colonized by an extensive population of *T. jerichonana* with tubes measuring up to 30-cm long; giant tube worms (*R. pachyptila*) were noticeably absent. Polyethylene markers 9, 10 and 16 (top numbered flap measures 30 × 15 cm) were deployed directly beside one another for subsequent recognition of the exact location of the area. By December 1993 (*b*), large numbers of giant tube worms (*R. pachyptila*) had settled and grown in the fissure. Marker 16 is in the lower left corner; markers 9 and 10 are obscured from view by the massive tube-worm colony. *b*, Video print produced by A. Giddings, E. Kristof, W. Lange and the *R/V Atlantis II/DSV Alvin* pilots and crew using the same systems and procedures used to produce the image illustrating the unusual sexual behaviour of deep-sea octopods reported in last week's issue on page 563.

observed spawning (as documented in other *R. pachyptila* populations)<sup>7</sup> on four separate occasions during the course of twelve *Alvin* dives during December 1993.

These documented growth rates (>85 cm per yr increase in tube length) of numerous individuals of *R. pachyptila* are the fastest reported to date for any species of vestimentiferan tube worm. The maximum earlier estimate<sup>5</sup>, which was based on indirect evidence, of the growth rate of this species was 14 cm yr<sup>-1</sup>. The maximum

- Grassle, J.F. *Science* **229**, 713-717 (1985).
- Lutz, R.A., Fritz, L.W. & Cerrato, R.M. *Deep-Sea Res.* **35**, 1793-1810 (1988).
- Tunnicliffe, V. *Oceanogr. Mar. Biol. A. Rev.* **29**, 319-407 (1991).
- Fustec, A., Desbruyeres, D. & Juniper, S.K. *Biol. Oceanogr.* **4**, 121-164 (1987).
- Roux, M. et al. *C. R. Acad. Sci., Ser. III* **308**, 121-127 (1989).
- Haymon, R.M. et al. *Earth planet. Sci. Lett.* **119**, 85-101 (1993).
- Van Dover, C.L. *Biol. Bull.* **186**, 134-135 (1994).
- Childress, J.J. & Fisher, C.R. *Oceanogr. Mar. Biol. A. Rev.* **30**, 337-441 (1992).
- Rhoads, D.C. & Lutz, R.A. *Skeletal Growth of Aquatic Organisms* (Plenum, New York, 1980).



growth rate of *T. jerichonana* reported herein ( $>30 \text{ cm yr}^{-1}$ ) also represents the highest growth rate reported for this species; the single published earlier estimate<sup>5</sup> which was based on a variety of assumptions, was  $9.2 \text{ cm yr}^{-1}$ . The only previous direct measurements of the growth of vestimentiferans were obtained from time-lapse camera studies of three individuals of *Ridgeia piscesae* inhabiting a hydrothermal vent field along the Juan de Fuca Ridge; incremental increases in tube length of the individuals ranged from 3.5 to 5.5 cm over 40 days (approximately  $30\text{--}50 \text{ cm yr}^{-1}$ )<sup>3</sup>.

Since the discovery of deep-sea, hydrothermal vents in 1977, there has been a growing body of evidence that rates of certain biological processes may proceed relatively rapidly in these unusual environments<sup>1-3,5,8</sup>. The observations reported here provide dramatic, unambiguous evidence that rates of colonization and growth of vestimentiferan tube worms in these hydrothermal systems are extremely rapid. The *R. pachyptila* growth rates, as measured as an increase in length per unit time, are not only the fastest

reported to date for any deep-sea organism, but they represent, to the best of our knowledge, the fastest rates of growth documented for any marine invertebrate<sup>9</sup>

**Richard A. Lutz, Timothy M. Shank**

*Institute of Marine and Coastal Sciences,  
Rutgers University,  
New Brunswick, New Jersey 08903, USA*

**Daniel J. Fornari**

*Department of Geology and Geophysics,  
Woods Hole Oceanographic Institution,  
Woods Hole, Massachusetts 02543, USA*

**Rachel M. Haymon**

*Department of Geological Sciences,  
University of California at Santa Barbara  
Santa Barbara, California 93106, USA*

**Marvin D. Lilley**

*School of Oceanography,  
University of Washington,  
Seattle, Washington 98195, USA*

**Karen L. Von Damm**

*Department of Earth Sciences,  
University of New Hampshire,  
Durham, New Hampshire 08324, USA*

**Daniel Desbruyeres**

*Deep Sea Environment Department  
IFREMER, Centre de Brest, BP70,  
29280 Plouzané, France*

ues may be overestimates of the expected cooling, because tropical sensitivities are less than the global IPCC sensitivities used here. Lapse rate considerations might allow for somewhat larger temperature depressions with height, but it seems unlikely that the snowline data can be reconciled with surface temperature changes less than about  $2.5\text{--}3.0^\circ\text{C}$ .

The above reasoning implies that it may be necessary to invoke relatively high climate sensitivities to explain ice age tropical SST changes, especially in regions far removed from the ice sheets. The implied sensitivity could be as high as  $4.5^\circ\text{C}$  for  $\text{CO}_2$  doubling. Such sensitivities are not necessarily out of line with the observed increase of global temperatures in the last century of only  $0.5\text{--}0.6^\circ\text{C}$  (ref. 15), because aerosols may be masking the effect of the  $\text{CO}_2$  perturbation<sup>16</sup>. But the ice-age data nevertheless suggest that, hidden within the apparent agreement between observed and computed global warming trends, the climate sensitivity may be quite high. If true, such a sensitivity could be dramatically manifested if aerosol production were to decrease<sup>17</sup>.

These calculations assume that the Guilderson *et al.* results are correct. Although some other data support such conclusions<sup>13,14,18</sup>, repeated efforts at further validating CLIMAP transfer functions<sup>19,20</sup> continue to reinforce the original CLIMAP conclusions. It now seems that clarification of ice age tropical SST changes is important not only for understanding past climates but also for better constraining predictions for the future.

**Thomas J. Crowley**

*Department of Oceanography,  
Texas A&M University,  
College Station, Texas 77843,  
USA*

## Pleistocene temperature changes

**SIR** — Guilderson *et al.*<sup>1</sup> provided geochemical evidence that tropical sea surface temperatures (SSTs) in the western Atlantic may have been  $5^\circ\text{C}$  colder during the last ice age. This conclusion appears to reconcile a discrepancy between tropical upland changes of  $5\text{--}6^\circ\text{C}$  (refs 2, 3) and CLIMAP<sup>4</sup> SST estimates of only a  $1\text{--}2^\circ\text{C}$  change. However, if the results in ref. 1 are substantiated by further work, they may create as big a climate problem as they solve, for it would then be necessary to explain why tropical SST changes were so large. To illustrate this point, SST decreases resulting from various changes in forcing can be estimated. For example, glacial-interglacial ice-age trace-gas changes have a radiative forcing of about  $2.6 \text{ W m}^{-2}$  (ref. 5). If we take the standard Intergovernmental Panel for Climate Change (IPCC) sensitivity estimates<sup>6</sup> for a  $\text{CO}_2$  doubling ( $1.5\text{--}4.5^\circ\text{C}$  for a  $4.2 \text{ W m}^{-2}$  increase) and apply those sensitivities to the ice age case, then the glacial-interglacial SST decrease due to trace gases should be  $1.0\text{--}2.9^\circ\text{C}$ . A substantial fraction of this range arises from uncertainties in cloud parameterizations<sup>7</sup>.

Model calculations indicate that ice sheets could cause an additional cooling of about  $1.5^\circ\text{C}$  in the tropical Atlantic<sup>8</sup>. If the value for the ice sheet response in this model is scaled to the IPCC range, then cooling from ice sheets would be  $1.0\text{--}2.9^\circ\text{C}$  in the tropical Atlantic. Although higher ice age dust loading could conceivably contribute another  $2\text{--}3^\circ\text{C}$  decrease<sup>9</sup>, the dust effect may have been overestimated

by a factor of two to ten.<sup>10</sup> Applying a median overestimate of a factor of five to Harvey's<sup>9</sup> dust calculation yields a revised cooling estimate from dust of about  $0.5^\circ\text{C}$ .

The combined potential SST response to the above changes is  $2.5\text{--}6.3^\circ\text{C}$  in the tropical Atlantic. This estimate does not consider additional uncertainties due to the ocean adjustment to changes in Atlantic deep- and intermediate water production<sup>11</sup>. Nor does it consider the possibility of a smaller temperature response to ice sheet changes if high dust levels significantly lowered the albedo of the ice sheets<sup>12</sup>. The potential effects of these changes on the tropical Atlantic is difficult to determine. The most that can be stated at present is that it is not inconceivable that the new geochemical estimates of Atlantic tropical SST changes are out of line with expectations. However, the results would seem to be more in line with sensitivity on the high side of the IPCC range.

Reconciling changes in the western tropical Pacific are more troublesome than for the western Atlantic. Snowline depressions are the same as they are elsewhere in the tropics<sup>2,3</sup> and preliminary geochemical results also suggest large late-glacial depressions in SSTs in this region<sup>13,14</sup>. Yet cooling from ice sheets, dust and ocean circulation changes should have been significantly less than in the western Atlantic. Thus, SST decreases in the western Pacific comparable to the Atlantic should primarily reflect ice age trace gas changes ( $1.0\text{--}2.9^\circ\text{C}$ ). These val-

1. Guilderson, T.P., Fairbanks, R.G. & Rubenstein, J.L. *Science* **263**, 663–665 (1994).
2. Webster, P.N. & Streeken, N. *Quat. Res.* **10**, 279–309 (1978).
3. Rind, D. & Peteet, D. *Quat. Res.* **24**, 1–22 (1985).
4. CLIMAP Project Members *Geol. Soc. Am. Map Chart Ser. MC-36* (1981).
5. Hansen, J., Lacis, A., Rued, R., Sato, M. & Wilson, H. *Res. Expl.* **9**, 143–158 (1993).
6. Mitchell, J.F.B., Manabe, S., Tokioka, T. & Meleshko, V. in *Climate Change: The IPCC Scientific Assessment*, (eds Houghton, J.T., Jenkins, G.J. & Ephraums, J.J.) 131–172 (Cambridge University Press, 1990).
7. Cess, R.D. *et al. Science* **245**, 513 (1989).
8. Manabe, S. & Broccoli, A.J. *J. geophys. Res.* **90**, 2167–2190 (1985).
9. Harvey, L.D.D. *Nature* **334**, 333–335 (1988).
10. Anderson, T.L. & Charlson, R.J. *Nature* **345**, 393 (1990).
11. Boyle, E.A. & Kellogg, L.D. *Nature* **330**, 35–40 (1987).
12. Warren, S.G. *Ann. Glaciol.* **5**, 177–179 (1984).
13. Beck, J.W. *et al. Science* **257**, 644–647 (1992).
14. Aharon, P. *et al. EOS (Trans. Am. Geophys. Un.)* **75**, 205 (1994).
15. Folland, C.K., Karl, T. & Vinnikov, K.Y.A. in *Climate Change: The IPCC Scientific Assessment*, (eds Houghton, J.T., Jenkins, G.J. & Ephraums, J.J.) 195–339 (Cambridge University Press, 1990).
16. Wigley, T.M.L. & Raper, S.C.P. *Nature* **357**, 293–300 (1992).
17. Wigley, T.M.L. *Nature* **349**, 503–506 (1991).
18. Piperno, D.R., Bush, M.B. & Colinvaux, P.A. *Quat. Res.* **33**, 108–116 (1990).
19. Prell, W.L. *US Dept. Energy Cont. Rep. Cont. No. DE-AC02-83ER60167* (1985).
20. Thunell, R., Anderson, D., Gellar, D. & Miao, Q. *Quat. Res.* **41**, 255–264 (1994).

# Now read on . . .

John C. Marshall

**The World on Paper.** By David R. Olson Cambridge University Press: 1994. Pp. 318. £17.95, \$24.95.

It cannot be in dispute that the transition from an illiterate to a literate culture entails a vast range of consequences over and above the ability to read this week's *Nature*. David Olson's fascinating (and infuriating) book advances a thesis that goes well beyond the (important) triviality that industrial or post-industrial societies would rapidly grind to a halt if a virus destroyed all written information (or our capacity to read it). His more grandiose claim is that we 'read' both the book of nature and the book of humanity *differently* after the world has been put on paper.

Whether one is reading between the lines or spelling out the message, it is clear that many of our metaphors derive from the written word. But is it really the case that "writing is largely responsible for bringing language into consciousness"? Or that Cartesian mentalism "was the by-product of a new understanding of what was in the text and what was contributed by a reader"? Or that the preliterate Huichol of Mexico believe that "corn is deer" is *literally* true?

With this latter example, Olson recklessly throws himself into some very muddy theological waters. The Huichol, it seems, regard both corn and deer as sacred. Furthermore, their "mythology" (a euphemism for the religions of 'primitive' peoples and destroyed cultures?) holds that "corn was once a deer". In such circumstances, it is difficult to see why their assertion that "corn is deer" is an instance of a "pre-logical mode of thought" characteristic of an illiterate society. Olson does, of course, realize that many literate peoples believe, for example, that wine turns into blood as a regular occurrence every Sunday morning. But this observation serves only to mire him yet deeper in the alleged reading-strategies of Maimonides, Aquinas and Luther.

In their view, Olson claims, "to read a text literally is to set the text not only in its context but also in terms of its putative writer and its putative reader". Quite so. But hearing an oral discourse correctly is similarly to appreciate the context in which it is said, and the character of the speaker and listener. Olson asserts that "a written transcription" in itself "captures only a privileged aspect of the utterance, namely 'what is said' and not 'how it is to be taken'". Well, yes . . . but does that make writing any different from speech? "Sincerity, seriousness and commitment are aspects of what is said which are not

represented by a script", he writes. True, they are indeed not represented in writing, but you need only listen to a politician to realize that neither are they reliably represented by what is spoken.

The central (and fatal) flaw in Olson's position arises from his premise that spoken language is a perspicuous and unambiguous expression of the speaker's mental state. This error leads him to contrast the (supposed) transparency of speech with the (purported) fact that



Is it or isn't it? *Ceci n'est pas une pipe* by René Magritte. (Reproduced from *The World on Paper*).

writing cannot readily represent the elocutionary force of an utterance. From this misconception, Olson argues that the reader must reflect on the written sentence in order to determine the speech act that it represents. The consciousness-raising involved therein then gives rise to the (post-mediaeval) distinction between the literal and the metaphorical. But once Olson's premise is seen to be false, the entire argument collapses.

I know only one story that (superficially) seems to confirm Olson's antithesis between speech and writing. It is told that, during a May Day parade, Stalin received a telegram from the exiled Trotsky. To the cheering crowds, Stalin read out: "You were right and I was wrong. You are the true heir of Lenin. I should apologize." And only Izzy the tailor, listening in the assembled throng, knew the telegram should really be read as: "You were right and I was *wrong*? You are the true heir of Lenin? I should apologize???!....." But that is why punctuation and italics were invented. Or, as the young Wittgenstein might have said, "anything that can be written at all can be written clearly".

Olson's conviction that the new 'literal'

reading of scripture entailed a more widespread cognitive change next leads him to argue that modern science adopted an equivalent form of reading the physical world: the emergence of astronomy from astrology, chemistry from alchemy, mathematics from numerology is akin to the difference between plain sense and oblique meaning in scriptural exegesis. Rashi's role in explicating the latter contrast is briefly mentioned but it is a pity that the much earlier clash between Karaism and Rabbanism is not. Olson's discussion of the conceptual origins of early modern science is entertainingly provocative; but it did not convince me that a causal relationship held between the revolutionary Protestant readings of scripture and how scientists interpreted God's other opus, the Book of Nature.

The line of argument is consistent, however, with Olson's belief in a very strong form of the hypothesis that literacy facilitates precision. "Traditional cultures", he writes, "treat alternative expressions of the same sense as being 'the same'; literate ones use the stricter criterion of verbatim repetition as 'the same'." But the claim that preliterate people (adults or children) do not make the distinction must surely be false. It is hard to imagine that traditional cultures do not have fixed forms of expression for significant rites and that a failure to use the correct

verbatim form would not render the rite null and void.

Similarly, one needs merely to repeat a story that begins "Once upon a time" a few times to a child before the introduction "Some time ago" will provoke howls of protest. Conversely, in our society, it would not be difficult for me to *paraphrase* the entirety of Olson's book and attempt to publish it as my own. I am, nonetheless, fairly sure that Olson would sue, and that all my protestations about the lack of "verbatim repetition" would cut little ice with the judge.

To be fair to Olson, I should point out that he scrupulously documents much of the evidence and argument *against* his thesis that "our modern conception of the world and our modern conception of ourselves are, we may say, by-products of the invention of a world on paper". It is just unfortunate that, on my reading of his text at least, the counter-evidence far outweighs the corroboration. □

John C. Marshall is in the Neuropsychology Unit, University Department of Clinical Neurology, Radcliffe Infirmary, Oxford OX2 6HE, UK.



## Uncommonly interesting

Paul Colinvaux

**Rarity.** By K. J. Gaston. *Chapman and Hall*: 1994. Pp. 205. £17.95 (pbk).

'RARE' is a term usually used by collectors to mean "I have got one and you haven't". Rarity is thus desirable, as reflected in one of its definitions in the *Oxford English Dictionary*: "thing valued as being rare". Biologists are not immune to this human foible, rejoicing in a rare bird or an unexpected species in the collecting net.

But for Gaston, in what is apparently the first monograph on biological rarity, 'rarity' merely means of low abundance or of small range, or both. The rare are ubiquitous, an inevitable part of any community. This abundance, as it were, of the rare is what makes rare species interesting. Why are certain species rare? Do ecosystems need them? Are they successful as a species? Is rarity a sign of impending extinction?

Frank Preston made a fundamental observation in his paper "The Commonness and Rarity of Species" (*Ecology* 29, 254; 1948) by showing that the distribution of relative abundance is log-normal. This implies underlying randomness. Subsequent attempts to find more subtle patterns of relative abundance have left few traces. Robert MacArthur's "broken stick" model, in which resources of a community are allocated as if by the random breaking of a stick into bits of unequal length, yields distributions of relative abundance that are themselves functions of the log-normal. Joel Cohen's variant of the broken stick theme, the "balls and buckets" model, in which population size is allocated by the random, sequential throwing of coloured balls into an army of buckets, yields a similar echo of the log-normal. No other models have fired the imaginations of ecologists as did the broken stick.

Gaston believes that interesting questions remain to be answered. He offers no new model or theory of rarity, nor is the history of the broken stick and its allies mentioned. Instead he begins a quest for properties of the rare that can be measured.

Definition is a major problem, and Gaston devotes his first chapter to it. He concludes that we have no alternative but to assign an arbitrary proportion of species present as rare, essentially the  $x$  per cent of species that have lowest abundance or smallest range. Five per cent of total abundance, biomass or range usually provides plenty of species in the 'rare' category.

With the rare so defined, Gaston takes a

population biologist's approach to measuring both the rare and the properties that might account for their rarity. Obstacles are formidable. The very rare almost cannot be counted, as when Eric Pianka noted that it takes a hundred person-days to find one specimen of a rare lizard. And to measure the area occupied by species with patchy distributions probably needs the use of fractals.

To investigate the causes of rarity by measurement, as the author wishes, without hypotheses of rarity to test, is more difficult still. Measuring arrays of environmental factors becomes uncomfortably reminiscent of older schools of plant-community analysis, where the aim was to explain plant distributions as consequences of physical parameters of environments alone.

Gaston describes several possible causes of rarity: endemism, individuals at range boundaries, vagrancy, the lows of fluctuating populations, restricted dispersal or establishment, as well as the pseudo-rarity that results from limits to our powers of detection. He dismisses the concept that some species remain rare because of restricted niche opportunities, because in the few examples where many resources have been measured, he is unable to find any correlation between resource and relative abundance.

As one who once tried to introduce Raymond Lindeman's concept of efficiency of energy transfer between trophic levels with an essay called "Why Big Fierce Animals are Rare", I remain convinced that tigers are rarer than sheep, and that I know why. But this monograph is about the rarity that remains after possible effects of food chains, functional niches, vagrants and sampling error are removed. This remainder is massive. Preston demonstrated the powerful influence of random process in bringing it about. Gaston suggests that experimental measures will find better answers. Perhaps. But it would be more fun to be given a new hypothesis of community-building over which we could argue with the passion ecologists once brought to the broken stick. □

Paul Colinvaux is in the Smithsonian Tropical Research Institute, PO Box 2072, Balboa, Panama.

### Correction

A line of text mysteriously disappeared from Michael Jacobs' review of *Environmental Politics and Greener Management International* in this year's New Journals supplement (*Nature* 371, 458; 1994). His full address is the Centre for the Study of Environmental Change, University of Lancaster, Lancaster LA1 4YN, UK.

It has also been brought to our attention that *Glaucois atlanticus*, the sea slug pictured on page 456 of the same issue, does not live on the sea bed, as stated. The organism is in fact an important inhabitant of the ocean surface film.

## Natural selections

Ray Percival

**Realism Rescued.** By Rome Harré, Jerrold Aronson and Eileen Cornell Way. *Open Court/Duckworth*: 1994. Pp. 203. \$42.95, £30 (hbk); \$18.95 (pbk).

How do you put both physicists and biologists on their guard? Answer: propound a philosophical theory that ignores Darwin's demolition of essentialism in species and brands any physicist who denies your theory of natural kinds as an anti-realist.

A traditional division in philosophy is between metaphysics (what sorts of things exist) and epistemology (what and how we know). Some think that the core of realism is the metaphysical assumption that there is a world independent of our minds. But this core assumption is sometimes clothed in other assumptions, such as theories of truth, truth-likeness, meaning and knowledge. Scornful of what they see as an unnecessary retreat from a fully clothed realism to the naked postulate of a mind-independent reality, Harré, Aronson and Way present a realism that also embraces truth and truth-likeness, as well as their own conception of scientific method and the structure of the world.

Informing their whole approach is a challenging view of scientific theories. Theories, they argue, are "essentially" models, or families of models, that constitute their "content". The idea that theories are sets of propositions is rejected. Well-constructed theories consist of a descriptive model, which portrays the phenomena, and an explanatory model, which portrays the unobservable substructure that causes the phenomena. Models are not simply dispensable aids to the construction and understanding of theories, as Duhem would have said. They expand a theory's explanatory power and help us to explain how theories can be continuously revised and extended to new phenomena.

Models represent type-hierarchies and type-hierarchies are pyramidal representations of the ordered hierarchy of natural kinds that make up the world. Borrowed from artificial intelligence, a type-hierarchy analysis, the authors say, is more faithful to our natural use of language, including metaphor and analogy. This is the authors' "naturalistic" approach. But defining realism in this way poses several problems.

In the authors' picture of the structure of the world, natural kinds are ordered in a hierarchy. Thus diamonds are a subtype of crystals, crystals are subtypes of molecular combinations, molecules are subtypes of combinations of atoms, atoms are subtypes of combinations of sub-

NATURE · VOL 371 · 20 OCTOBER 1994

atomic particles. In the 1960s, however, some people held an interesting non-atomistic (holistic) view of matter that pictured each proton and neutron as made of one another. This was called a "bootstrap theory" after Baron von Munchhausen, who claimed to have lifted himself up by his own bootstraps. In this view, the hierarchy cannot be continued because the constituents of the hadrons are no more elementary than the given hadrons — they would be subtypes of each other. This theory is now refuted, but some might quibble at the suggestion that entertaining such a theory at the time was "anti-realist". Defining realism in terms of the details of a particular ontology means that the critics of this particular ontology are anti-realists by definition!

Darwin's theory undermines talk of natural kinds. The ancestry of both a modern human and an earthworm can be traced back through intermediate species to a common virus-like ancestor. How could this be if the virus-like ancestor was a natural kind, that is, had an essence? Darwin's answer was that the steps up to and over the "boundary" between one species and another are slight.

The authors cover in a very readable way many problems, including truth, truth-likeness, metaphor and analogy, similarity, contrary-to-fact-conditionals, and primary and secondary qualities. Truth itself is defined in terms of their view of the structure of the world. Instead of defining realism in terms of a propositional notion of truth as correspondence, truth is defined in terms of realism (the ordering of natural kinds). A model is true if the chunk of the type hierarchy it picks out is identical to the section of the ordered relationships between natural kinds in the world. A model is close to the

truth if there is a similarity between these two type-hierarchies.

My reservations about the potential of the project are prompted by the authors' acceptance of a group of refuted doctrines, including naturalism, induction, essentialism and antilogicism. Naturalism is the idea that we should assess philosophical doctrines in terms of what psychologists or artificial-intelligence specialists find out about the way we form concepts, reason and so on. A naturalistic approach to the philosophy of science, however, scorns a real distinction between a descriptive analysis of science and a prescriptive (methodological) analysis. There are several dangers with this. Appraisal of method and theory is liable to be subject to mere fashion. If the 'best method' is understood tacitly as what scientists happen to do and this happens to be unadventurous, then the creation and acceptance of theories that challenge the orthodoxy are impeded. Naturalism also has a tendency to confuse the cognitive (psychological) mechanisms that produce a theory with the nature of the theory considered as an objective product with autonomous properties and its appraisal, just as a proud carpenter might overvalue his product because of the many long hours he spent fashioning it.

There is much to disagree with here, but the systematic presentation of the model view of scientific theories will produce useful debate. □

*Ray Percival is at 70 Hillview Court, Astley Bridge, Bolton BL1 8NU, UK, and is organizer and chairman of the Annual Conference on the Philosophy of Sir Karl Popper and associate editor of the Popper Newsletter and Journal of Social and Evolutionary Systems.*

## No stones left unturned

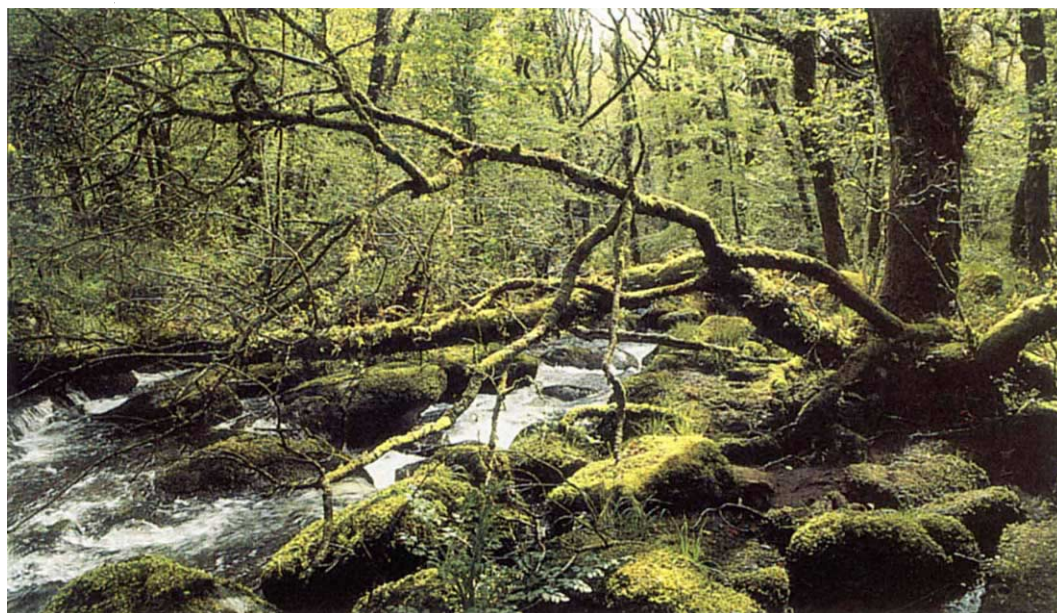
*James A. Tyburczy*

**Introduction to the Physics of Rocks.** By Yves Guéguen and Victor Palciauskas. Princeton University Press: 1994. Pp. 294. \$49.50, £35.

A GEOPHYSICIST studies Earth's subsurface by measuring bulk properties such as seismic velocity, electrical conductivity, thermal conductivity or magnetization from the surface or in a borehole. Then, using the tools of rock physics, he or she extracts from these measurements information about the physical state of the rocks beneath the surface, such as rock type, porosity, permeability and fluid saturation. In this way, one can explore for hydrocarbons, minerals and geothermal resources, and study natural and man-made hazards such as earthquake faults, groundwater pollution and nuclear storage schemes.

Crustal rocks are heterogeneous mixtures of minerals, generally porous and/or fractured and commonly partly to fully saturated by fluids such as water and hydrocarbons. So the tools of the rock physicist include not only elasticity theory, continuum mechanics, rock deformation and fracture mechanics (the sum of which is often referred to as 'rock mechanics'), but also theories of heterogeneous and porous media, fluid flow and percolation theory, surface chemistry, electric and dielectric theory, heat transfer and the theory of magnetic materials.

The text fills a unique and previously



**BOVEY Valley Woodlands, Devon, one of the 140 wildlife sanctuaries featured in England's National Nature Reserves** by Peter Marren (English Nature/T&AD Poyser, £20). With line drawings and over 60 colour plates, the book provides a comprehensive look at the background and philosophy of the reserves and the role they play in conservation and research.



vacant niche by providing a comprehensive, up-to-date and accessible introduction to Earth-science applications of modern theories in rock physics. The authors lay the groundwork by developing the basic description of the microstructures of porous and heterogeneous rocks. They discuss developments in chemistry and physics applicable to rock physics, such as surface adsorption of ions, effective medium theories for calculating bulk properties of heterogeneous and porous media, and fractal descriptions of fractures and pore surfaces, and then use these concepts to discuss acoustic properties, electrical conductivity, dielectric properties, thermal conductivity and magnetic properties of bulk rocks, particularly as influenced by fluids.

Two features set the book apart from others. First, it is comprehensive. The chapters on electrical conductivity, dielectric properties, thermal conductivity and magnetic properties, as well as the more traditional sections on the mechanical properties of rocks and flow through porous media, make this the only text to cover virtually all of the physical prop-

erties that concern rock physicists. Second, the text is well integrated, with concepts introduced early on in one context reapplied later in another. For example, effective medium theory is introduced in the discussion of elastic properties of heterogeneous media, and then reappears in chapters covering electrical, dielectric and thermal conductivity. Students will find this continuity helpful and enlightening.

The book is a translation of a volume originally published in French, *Introduction à la Physique des Roches* (Hermann Éditeurs des Sciences et des Arts, Paris, 1992). It has abundant references to recent work (up to about 1990), a feature that will enable readers to have good access to the current literature.

It will be useful for a rigorous graduate course on the physical properties of rocks, as well as being valuable as a reference for professionals interested in the subject. □

James A. Tyburczy is in the Department of Geology, Arizona State University, Box 871404, Tempe, Arizona 85287-1404, USA.

Coincidentally, Celia Heyes (*Animal Behaviour* 47, 909; 1994) has argued (with some justification, in my view) that owing to the lack of proper controls, the Gallup mark-test does not in fact show evidence of mirror-use. The obvious control procedure of, say, simultaneously sham-marking the chin when paint is applied to the forehead, has been neglected. Ironically, something similar to this procedure is here reported with great effect in a gorilla, the one great ape species hitherto dismissed as lacking self-recognition. Together with spontaneous mirror use, this evidence is compelling — much more so than that presented for monkeys and dolphins. Although a monkey seems to see another monkey in its reflection, it is beginning to look as if, in the case of all four genera, a great ape may in some sense see itself (although not all do, and Daniel Povinelli suggests an interesting explanation).

But just because apes can see their self reflected does not necessarily imply the other sense — that they self-reflect. When it comes to the psychological meaning of these patterns, nobody knows. The book does show that theory is developing, some at least in the direction of testable predictions. The integration of findings from developmental, abnormal and comparative psychology is already used in this spirit by several authors. Scarcely a stone is left unturned in laying out the theoretical options, and particular credit is due both to Parker and to Mitchell for their efforts to this end. □

Andrew Whiten is in the School of Psychology, University of St Andrews, St Andrews, Fife KY16 9JU, UK.

## Solving the monkey puzzle

Andrew Whiten

**Self-Awareness in Animals and Humans.** Edited by Sue Taylor Parker, Robert W. Mitchell and Maria L. Boccia. Cambridge University Press: 1994. Pp. 442. £40, \$59.95.

WHAT do you see when you look in the mirror? Why, your self, of course, you will be thinking to yourself. But here are two quite different senses of 'self' — the self we see reflected and the self who reflects. Our sense of self is so central to us that questions about the emergence of such forms of self-awareness in evolution or development could hardly be more fundamental. And if answers can be provided, they may well be seen as relevant to our codes of morality concerning, for example, the ethical treatment of animals.

One of the most influential steps towards a scientific approach to such apparently ineffable issues came in the 1970s, when Gordon Gallup showed that chimpanzees and orangutans would not only use mirrors to inspect themselves, but would wipe off marks applied to their faces under anaesthetic; yet monkeys failed to do these things. Gallup suggested that the apes alone, like humans, were showing an awareness of self, from which further capabilities could arise, such as intentional deception and empathy.

Nearly all of the 16 'animal' chapters in this book are about these responses to mirrors: to accept the title as most people

would read it, one has to buy Gallup's generous interpretation. But if one accepts this focus, the volume offers a comprehensive and indispensable survey of the state of the art, presented by today's principal workers and including interesting extensions of the approach to other taxa.



A NOSE for self-reflection? A chimpanzee in one of Gallup's experiments (see above) explores the uses of a mirror. The photograph appears in *The Animal Mind* by J. L. and C. G. Gould (W. H. Freeman, \$32.95), which will be reviewed by Marian Stamp Dawkins in a future issue.

# Evidence from gravity and topography data for folding of Tibet

Yu Jin<sup>\*</sup>, Marcia K. McNutt<sup>\*</sup> & Yongsheng Zhu<sup>†</sup>

<sup>\*</sup> Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

<sup>†</sup> Department of Geology, Ocean University of Qingdao, 266003 Qingdao, China

**Bouguer gravity and topography data from Tibet suggest that the surface of the plateau and the subsurface density interfaces are warped into two series of ridges and troughs trending parallel to the collision zone with wavelengths of 150 and 500 km. These folds are superimposed on an overall state of isostatic compensation owing to crustal thickening. Such folding is predicted from models of compressional shortening of a rheologically layered plate. The results thus suggest the presence of a weak decoupling zone between the Tibetan crust and upper mantle.**

THE Tibetan Plateau is the most conspicuous feature of continental topography on Earth, and considerable effort has been devoted to understanding the cause and timing of its uplift for the purpose of understanding its influence on the tectonics of Asia<sup>1</sup>, the development of the Indian monsoons<sup>2</sup> and long-term climate patterns<sup>3</sup>. All models for the formation of this broad plateau begin with the assumption that the crustal thickness beneath the plateau has been approximately doubled in response to the collision of India with Asia, beginning about 45 million years ago<sup>4</sup>, but differ in the details of how this doubling is accomplished. One of the earliest proposals<sup>5,6</sup> involved simple underplating of the Tibetan crust by the Indian crust. A more recent update of this model conserves mass beneath the crust by proposing subduction of the Asian mantle lithosphere beneath northern Tibet, and involves more realistic compressional shortening as well as overthrusting of Indian crust<sup>7</sup> (Fig. 1a). The "rigid indenter" models<sup>8–10</sup> assume that the entire continental lithosphere behaves as a viscous fluid that has been shortened by a factor of two, and thus thickened, by the northward advance of India (Fig. 1b). This model predicts thickening of the lithospheric upper mantle as well as the crust, although in a variation on this theme by England and Houseman<sup>11</sup>, some portion of the thickened and dense upper mantle detached within the past 11 Myr, leading to rapid uplift of the plateau. Alternatively, Zhao and Morgan<sup>12</sup> have proposed injection of the Indian crust into the weak Tibetan lower crust, leading to plateau uplift in the manner of a hydraulic jack. The fate of the Indian mantle lithosphere is unspecified; one might assume that, stripped of its buoyant continental crust, the Indian upper mantle would subduct into the asthenosphere.

The models discussed above differ in their predictions for both the strain field and the strength profile of the Tibetan block. For example, the indenter model (Fig. 1b) requires high (~50%) strain and low strength at all depths. In contrast, the Asian subduction and crustal injection models (Fig. 1a and c, respectively) require low strength only in the decoupling zone in the midcrust (where Indian crust is either underthrust or injected) and subduction rather than north-south shortening in the upper mantle beneath Tibet. These latter two models do differ in whether the upper mantle beneath Tibet is Indian upper mantle (Fig. 1a) or the original Tibetan upper mantle (Fig. 1b), but it may be difficult to distinguish the two possibilities on the basis of rheological inferences because the cold, stiff upper mantle beneath the Indian plate<sup>13,14</sup> could be weakened and reheated in the process of underthrusting such that it is indistinguishable from the younger upper mantle beneath the Mesozoic Tibetan block.

Here we present high-resolution gravity data from the Tibetan Plateau that allow us to estimate both its strain and its strength profile. The development of two scales of folds in the surface and subsurface at wavelengths of 150 and 500 km demonstrate that the crust has clearly shortened. However, the coherence between gravity and topography indicates that the upper mantle is relatively rigid beneath Tibet and resists compression, thus lending support to the subduction scenario<sup>7</sup>. Based on both the observation of the development of two wavelengths of folding and results from the experimental deformation of rocks, we propose that Tibet contains two distinct strong regions: one in the upper crust and the other in the uppermost mantle, separated by a lower-crustal decoupling layer.

## Bouguer gravity and topography data

The Bouguer gravity data and land elevation information we use was compiled by Ocean University of Qingdao, People's Republic of China. The original point gravity measurements were based on the Potsdam standard within the geographical reference of the Beijing coordinate system and the Yellow Sea elevation system. Data were reduced using the Helmert normal gravity formula for gravity on the spheroid and converted to free air anomalies. The Bouguer gravity field was calculated using a topographic density of 2,670 kg m<sup>-3</sup>, and terrain corrections were applied out to a distance of 166.7 km (ref. 15). Through a scientific collaboration agreement between the Ocean University of Qingdao and the Massachusetts Institute of Technology, the gravity and topography data were interpolated for us onto a 10-km grid. The original point measurements were uniformly distributed over the plateau, but the lower limit on the wavelengths resolved is closer to 25 than 10 km, except in a few well-sampled areas. The cumulative error in the Bouguer gravity field, including errors in elevation, is estimated to be 1.5 mGal (ref. 15), although uncertainty in the Bouguer reduction density is not included in this value. The validity of the topographic data was verified by E. Fielding at Cornell University by comparing our values over a 2° × 2° area with an independent, even higher-resolution, digital elevation model<sup>16</sup>. Although there is a small discrepancy in the registration of the two topographic data sets, the average differences are less than 10 m.

As shown in Fig. 2, the land elevation in our study area consists of east-west trending undulations about a mean plateau elevation of 5 km. Except near the edges of the plateau, the relief rarely departs by more than 1 km from the mean value. This uniform elevation despite the large amount of crustal thickening beneath the plateau has been one of the principal arguments in favour of a weak crust that flows laterally to even out any thick-



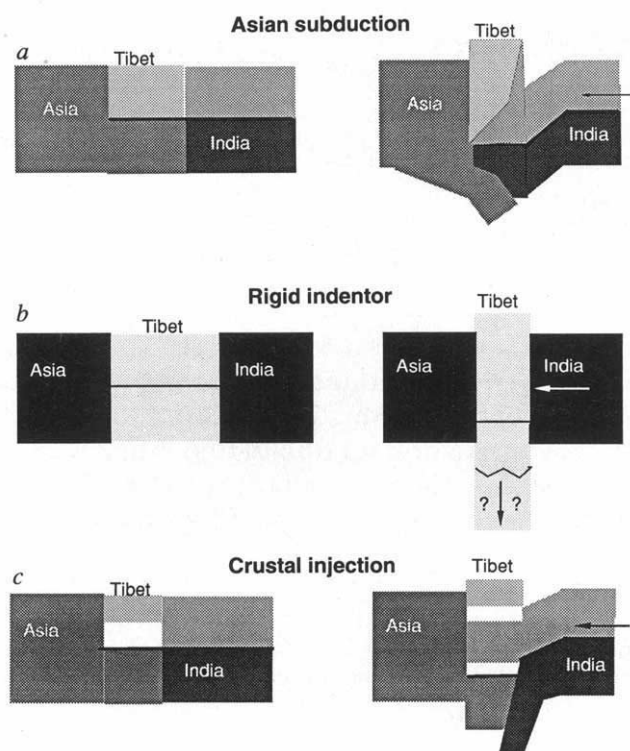


FIG. 1 Three models for the formation of the Tibetan Plateau. Asia, Tibet and India are each represented as blocks in cross-section. The left and right panels for each model show the crustal structure before (left) and after (right) collision with India. Darker shading indicates relatively more rigid blocks. The black line shows the location of the crust-mantle boundary for Tibet and India. The Asian subduction model (a) thickens the Tibetan crust by thrusting Indian crust and mantle beneath the Tibetan crust. The Asian upper mantle is subducted. The rigid indenter model (b) thickens the Tibetan crust and upper mantle by shortening a uniformly weak block. The crustal injection model (c) intrudes the Indian crust into the soft Tibetan lower crust and subducts Indian upper mantle.

ness irregularities imposed during the crustal thickening process<sup>17</sup>.

Modelling of the gravity and topography data measured along cross-sections perpendicular to the Himalayan thrust front (Fig. 3) confirms that, to first order, the Bouguer data are fitted by local (Airy) isostatic compensation (solid lines, Fig. 3). Therefore, the overall high elevation of the plateau is buoyed up by a thickened crust with no apparent contribution from any thermal anomalies in the upper mantle. The most notable departures from the predicted Airy model occur just north and south of the plateau's margins, where the more negative Bouguer anomalies reflect the development of foredeep basins on bordering plates with appreciable flexural strength. In addition, there is noticeable signal in the Bouguer gravity at shorter wavelengths undulating across the entire plateau this is not predicted by local compensation, indicating that there must be uncompensated topographic features and/or buried mass anomalies with no corresponding topographic expression. This simple observation suggests that some strong zone must exist within or beneath the crust to support short-wavelength topographic features and prevent buried loads from achieving local isostatic balance by deforming the surface.

### Coherence between gravity and topography

Forsyth<sup>18</sup> pioneered the technique of using the coherence between Bouguer gravity and topography to estimate the effective elastic thickness in continental regions containing both surface and subsurface loads on the lithosphere. If  $\Delta G$  and  $H$  are the Fourier transforms of the Bouguer gravity field and topography,

respectively, then coherence  $\gamma^2$  is defined as

$$\gamma^2 = \frac{\langle \Delta G \cdot H^* \rangle \langle \Delta G \cdot H^* \rangle}{\langle H \cdot H^* \rangle \langle \Delta G \cdot \Delta G^* \rangle} \quad (1)$$

in which the asterisk denotes complex conjugation and the angle brackets represent averaging in bands of constant wavenumber modulus  $k$ . The essence of Forsyth's argument is that if a plate has no elastic strength, then all surface loads will produce corresponding crustal thickness variations at the Moho to obtain local isostatic equilibrium, and all imposed crustal thickness variations will produce corresponding topography. The Bouguer gravity anomaly, which is sensitive to those undulations at the Moho, will correlate at all wavelengths with the topography and the coherence will be uniformly high. In the other extreme, if the lithosphere is perfectly rigid, surface loads will produce no flexure at the Moho, and loads buried within the plate will produce no topography. The topography will be a perfect measure of the surface load and the Bouguer gravity will reflect only the subsurface load. If the surface and subsurface loads are uncorrelated, the coherence between Bouguer gravity and topography will be small at all wavelengths. Any realistic case for a plate with finite elastic strength will fall somewhere between these two extremes, and the wavelength at which the coherence drops from values approaching unity at long wavelengths to values approaching zero at short wavelengths is a measure of the rigidity of the elastic plate.

For the case in which the subsurface load is applied at the Moho, the theoretical coherence for an elastic plate is given by

$$\gamma^2 = \frac{(1+f)^2}{(1+Y^{-2}f)(1+Y^2f)} \quad (2)$$

$$Y = \left[ 1 + \frac{(2\pi k)^4 D}{\Delta \rho g} \right]^{-1} \quad (3)$$

where  $f$  is the ratio of the power in the subsurface load spectrum to that of the surface load,  $D$  is the flexural rigidity of the plate,  $\Delta \rho$  is the density contrast at the Moho, and  $g$  is the acceleration due to gravity.  $D$  is related to the elastic thickness  $T_e$  by  $D = ET_e^3 / 12(1-\nu^2)$  where  $E$  is Young's modulus ( $10^{11}$  N m<sup>-2</sup>) and  $\nu$  is Poisson's ratio (0.25).

Figure 4 compares the observed coherence calculated for the entire Tibetan Plateau (filled circles) to the predicted coherence if the surface and subsurface loads are supported elastically (solid lines). The drop in the coherence from high values at long wavelengths to low values at short wavelengths begins at wavelengths greater than 500 km, as would be predicted if the elastic plate supporting the plateau's topography is at least 40 km thick. However, the observed coherence does not drop to zero as rapidly as predicted by the theoretical models, and displays an anomalous peak at wavelengths of 100–200 km. Bechtel *et al.*<sup>19</sup> have suggested that this behaviour might be caused by averaging together in the analysis geological provinces with different elastic rigidities, for example, in this case, between 10 and 40 km elastic thickness. But this does not seem to be the explanation here; even when we subdivide the plateau into three sub-regions (Fig. 2 inset) representing geologically homogeneous provinces, the observed coherence points (diamonds, triangles, and squares in Fig. 4) are similar for each one. Clearly the support for the plateau is not simply described by compensation of uncorrelated loads on and within a single elastic lithosphere.

Further investigation of the gravity spectrum and its relationship to the topography spectrum suggests that one reason for the failure of the simple model to fit the coherence data is that Tibet is layered in both its density and its rheological structure. The slope of the logarithm of the gravity power spectrum<sup>13,20</sup> (Fig. 5a) implies that at wavelengths greater than 100 km, the Bouguer gravity signal is caused by undulations of density interfaces at about 13 km (a midcrustal discontinuity we refer to as the "Conrad") and at about 60 km (the Moho). The two-

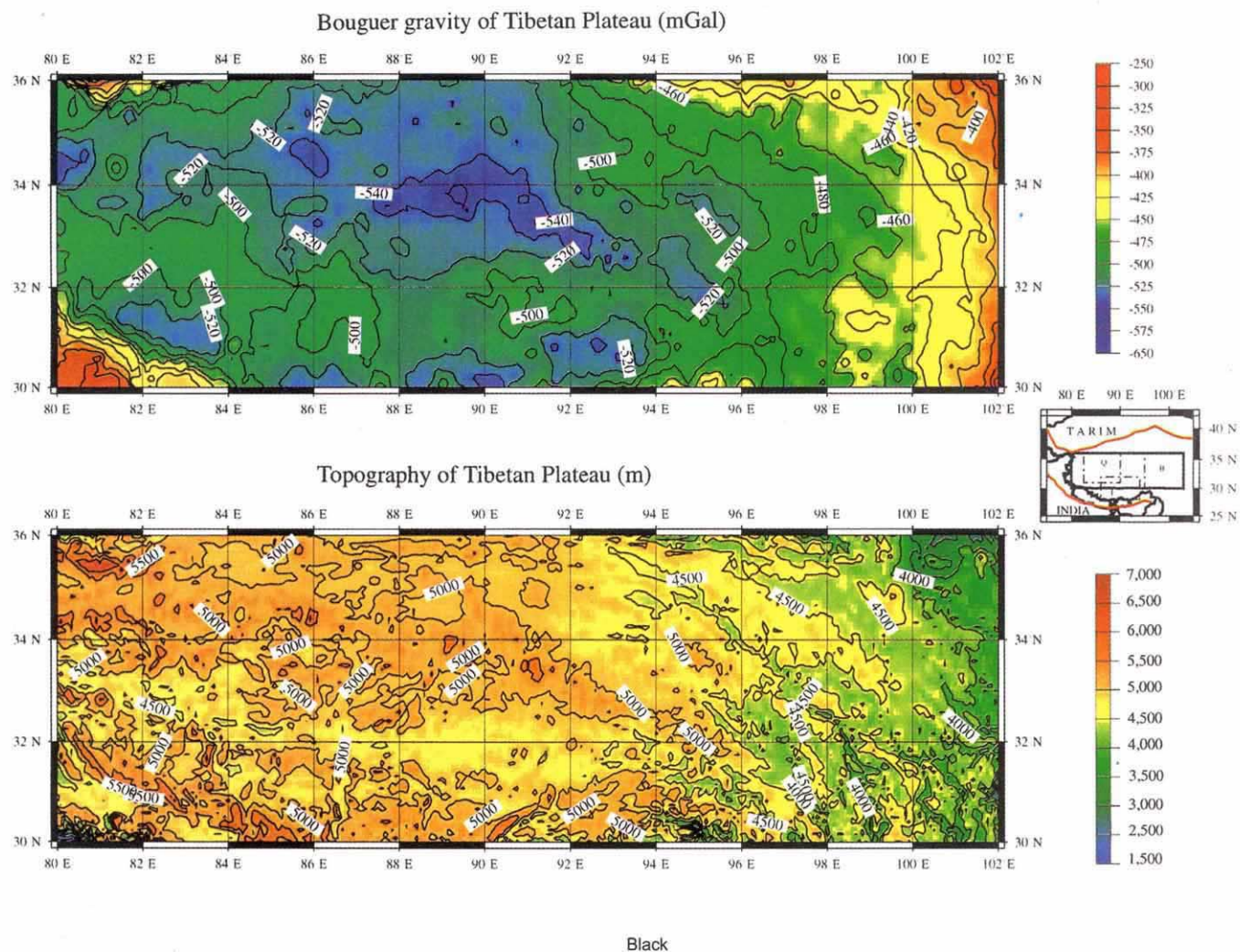
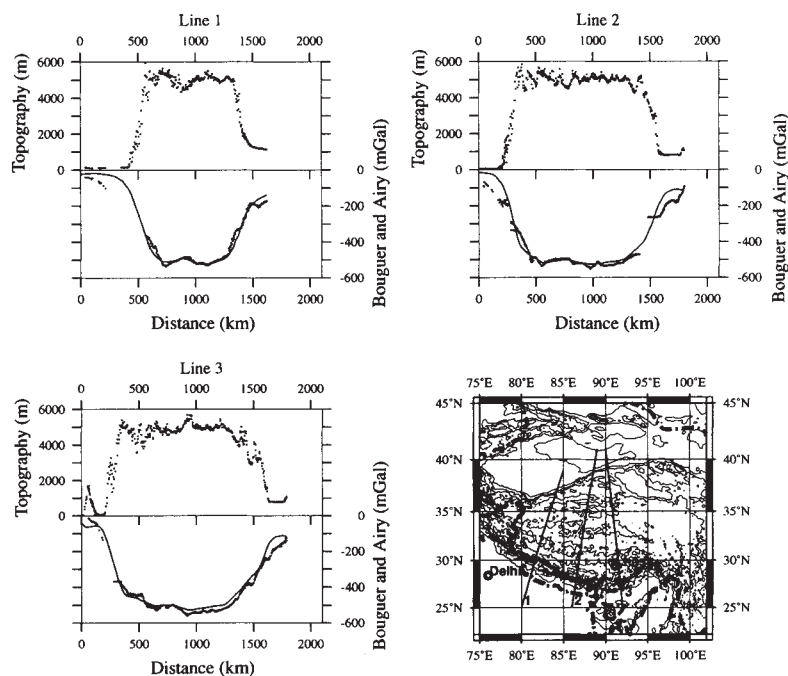


FIG. 3 Graph panels, profiles of topography and Bouguer gravity (dots) along three cross-sections oriented perpendicular to the Himalayan thrust front spanning the plateau from the Indian plate to the south (left) to the Tarim Basin on the north (right). Solid lines show the predicted Bouguer gravity assuming an Airy model of local isostatic compensation (no elastic rigidity) with a crustal thickness of 60 km, crustal density of  $2,670 \text{ kg m}^{-3}$  and mantle density of  $3,300 \text{ kg m}^{-3}$ . Map panel, locations of the cross-sections superimposed on a topographic map of the region around Tibet.





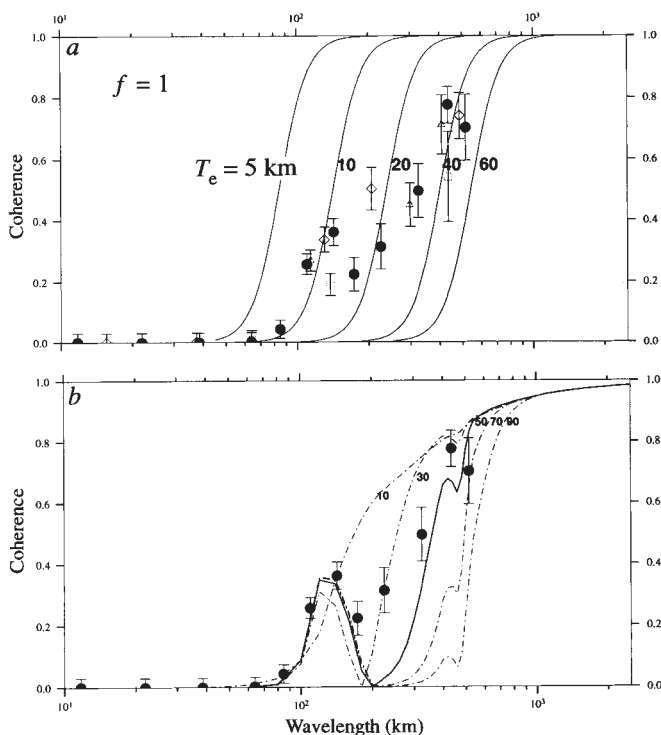
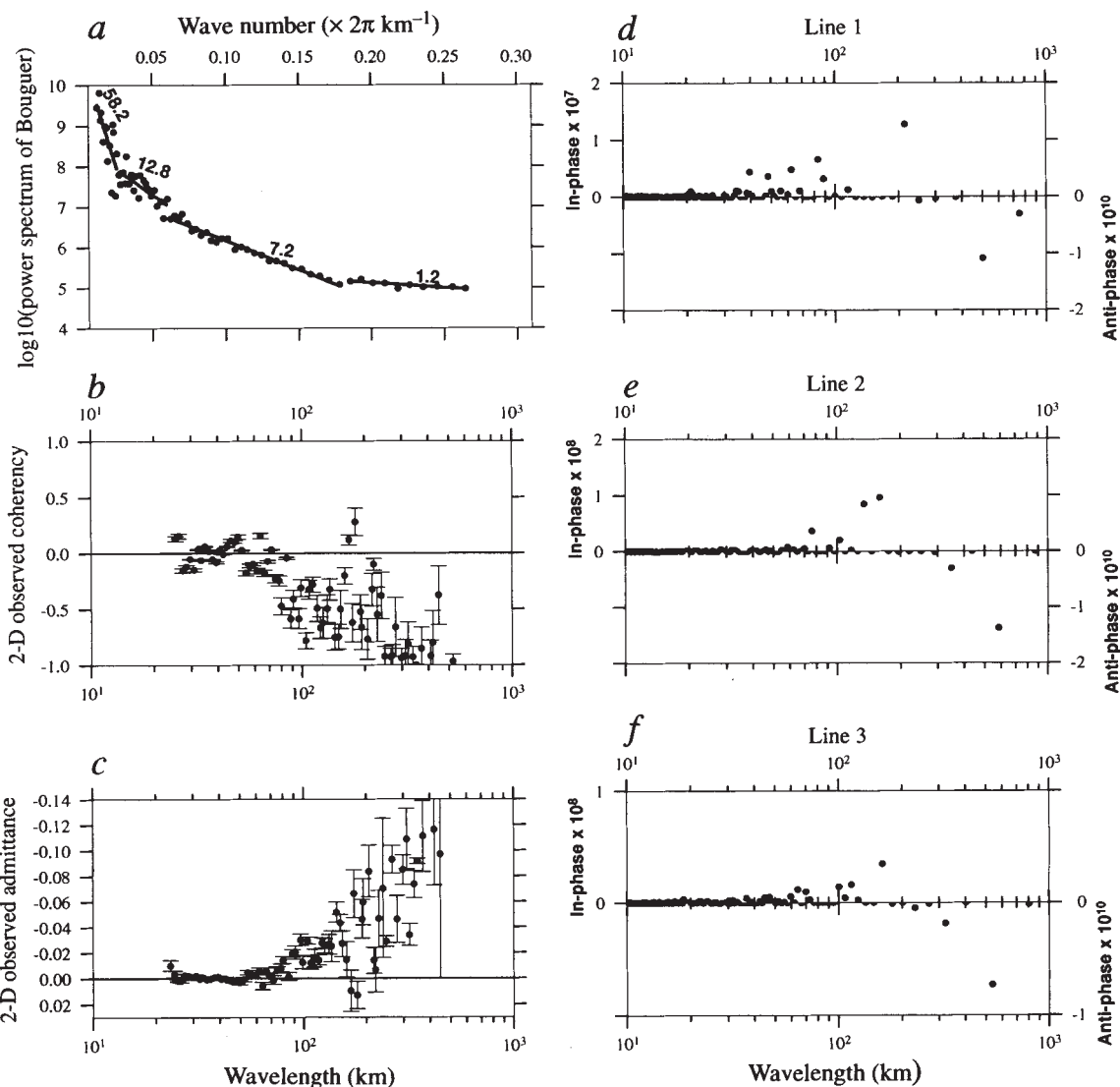


FIG. 4 *a*, Comparison of the observed coherence calculated for the Tibetan Plateau with that corresponding to theoretical elastic plate models as a function of wavelength in kilometres. The solid circles give the coherence for the entire Tibetan Plateau, whereas the triangles, squares and diamonds represent coherence for the Qiangtang, Lhasa and Bayanhar sub-regions (Fig. 2 inset), respectively. The solid lines correspond to theoretical curves for five different values of the elastic plate thickness,  $T_e$ , assuming uncorrelated surface and subsurface loads and a single subsurface density discontinuity at the Moho. For the case of a subsurface loading only at the Moho, the theoretical coherence is relatively insensitive to the value for  $f$ , the ratio of subsurface to surface loads, here assumed to be 1. *b*, Comparison of the observed coherence data for the entire Tibetan Plateau with theoretical coherence as a function of wavelength in kilometres allowing for the correlation of loads on the surface and in the subsurface. Subsurface loading is allowed at the Conrad discontinuity at 13 km and at the Moho at 60 km. Folding of a rheologically layered plate is simulated by allowing a 60% positive correlation between surface and Conrad loads in the two wavelength bands for folding at 150 and 500 km. A 60% negative correlation between surface and Moho loads is used in the 500-km band. For a layered plate, the curves are more sensitive to  $f$ . For these curves, we assumed  $f=3$  for both the Conrad and Moho loads.

FIG. 5 *a*,  $\log_{10}$  of the power in the Bouguer gravity spectrum as a function of wavenumber ( $2\pi/\text{wavelength}$ ). Straight lines fitted piecewise to the slope of the spectrum provide an estimate of the depth to the density discontinuity responsible for that portion of the gravity spectrum. Numbers on straight line segments represent the average depth to the density discontinuity presumably responsible for that portion of the gravity spectrum. *b*, Coherency of Bouguer gravity and topography for Tibet. The trend from coherency values near zero at short wavelengths to negative values at long wavelengths as expected, for isostatic compensation is interrupted by positive values at 150–200 km. *c*, Admittance of Bouguer gravity and topography for Tibet. *d–h*, One-dimensional cross-correlation of topography and Bouguer gravity for the profiles across the plateau in Fig. 3. Note the large difference in scales for the positive, in-phase, correlation as opposed to the negative, out-of-phase, correlation necessary to highlight the weak positive correlation on all profiles between 150 and 200 km. The prominent negative signal on each profile occurs at 500–700 km wavelength.



dimensional coherency  $\gamma$  (Fig. 5b) and admittance  $Q$  (Fig. 5c) between gravity and topography, defined as

$$\gamma = \frac{\langle \Delta G \cdot H^* \rangle}{\langle H \cdot H^* \rangle^{1/2} \langle \Delta G \cdot \Delta G^* \rangle^{1/2}} \quad (4)$$

$$Q = \frac{\langle \Delta G \cdot H^* \rangle}{\langle H \cdot H^* \rangle} \quad (5)$$

both show an unexpected positive correlation between Bouguer gravity and topography at wavelengths of 150–200 km which is not predicted by any conventional model of isostatic compensation. One-dimensional cross-spectra  $\langle \Delta G \cdot H^* \rangle$  (Fig. 5d–f) computed from the profiles in Fig. 3 similarly display a weak positive peak in the correlation at 150–200 km, as well as a strong negative peak at 500–700 km. We interpret these peaks in the cross-correlation spectrum as evidence for the development of folds<sup>21</sup> in the Tibetan plate by the collision with India, just as the sea floor south of the Bay of Bengal has been folded by this collision<sup>22</sup>, with the two distinct wavelengths arising from the expected multi-layered rheology of continental crust<sup>23–25</sup>.

### Folding of Tibet

Figure 6 shows a strength envelope as inspired by experiments on the deformation of rocks extrapolated to realistic geological pressures, temperatures and strain rates. Strength in the upper part of the plate is controlled by frictional sliding along pre-existing faults. The crust becomes increasingly strong with increasing overburden pressure down to depths of about 15 km, at which point ductile creep of diabase becomes the dominant mechanism of failure. Between depths of 25–35 km (depending on the local geotherm) and the base of the crust, a weak ductile channel should exist. In the case of Tibet where the crust is 60-km thick, the ductile channel should be ~30-km thick and well developed. In the uppermost mantle, another relatively strong zone should exist based on the higher strength of olivine. When subjected to the moderate stresses from vertically-directed surface and subsurface loads, the overall mechanical behaviour of the two strong regions separate by a ductile channel has an effective elastic thickness given by  $(T_c(\text{upper crust})^3 + T_c(\text{upper mantle})^3)^{1/3}$ , which in this case is a little more than the effective elastic thickness of the stronger of the two layers.

However, if subjected to end loads of the order of several hundred megapascals, as might be produced by the collision of India with Tibet, the strength of the upper-crustal layer will be exceeded everywhere. It will fold as a plastic/viscous layer. The stronger layer in the upper mantle will also fold, presumably in an elastic/plastic manner<sup>26</sup>. Zuber<sup>27</sup> has shown that the deformation of two strong layers separated by a weak ductile zone subjected to horizontally-directed compression will be characterized by the growth of folds at two dominant wavelengths, which might explain the signals at 150 and 500 km in the Bouguer gravity data.

We can use the coherence data to test the hypothesis that the correlation between Bouguer gravity and topography at 150-km and 500-km wavelengths is caused by the folding of Tibet. The theoretical curves in Fig. 4b correspond to the coherence predicted for elastic plates loaded on density interfaces at the surface, a midcrustal discontinuity at 13 km, and at the Moho, as suggested by the gravity spectrum in Fig. 5. The loads at these three interfaces are assumed to be uncorrelated except in narrow wavelength bands centred around 150 and 500 km, where the coherency between the surface and subsurface loads is set at 60% to simulate the effect of folding. In keeping with the folding model shown in Fig. 6, we require that the correlations between the surface and midcrustal loads be positive in both wavelength bands. The correlation between the surface and Moho loads is only non-zero in the 500-km band, where the sign is negative. For large values of the ratio of midcrustal to surface loads ( $f \approx 3$ ), the best-fitting elastic plate thickness is approximately 50 km (Fig. 4b). Smaller values of  $f$  lead to smaller values of

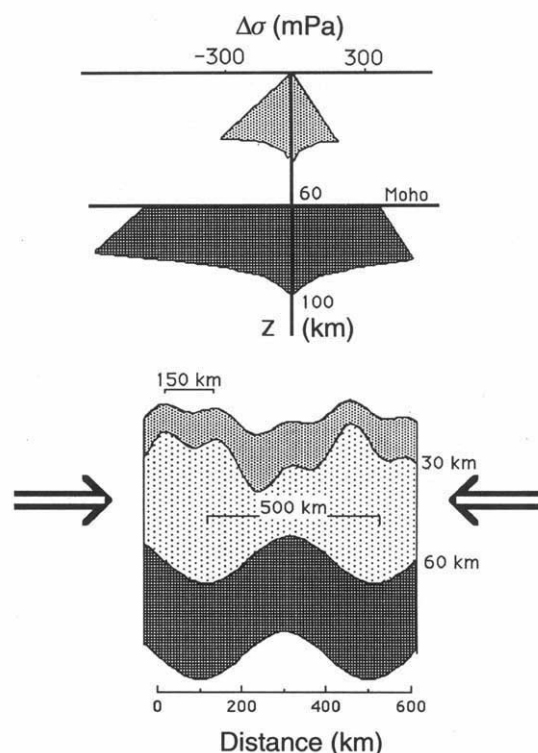


FIG. 6 Top, strength envelope for the crust and lithospheric upper mantle<sup>25</sup>.  $\Delta\sigma$  is the difference between the maximum horizontal and vertical stresses in the plate and  $Z$  is depth. Strength in the upper part of the crust and uppermost mantle is limited by frictional sliding on faults and increases with overburden pressure. Thermally-activated creep relaxes elastic stresses in the lower crust and below the uppermost mantle. If the stress at any depth in the deformed lithosphere lies within this strength envelope, the stresses will be supported elastically (these regions are denoted by light grey shading for the upper crust and dark grey for the upper mantle). Stresses exceeding the elastic limit will cause failure by frictional sliding or ductile flow. Bottom, cross-section of lithosphere under compression that develops two scales of folding in a rheologically layered plate. The upper plate deforms in a viscous or plastic mode and contains the Conrad discontinuity. Warping of this discontinuity parallel to the surface leads to a Bouguer gravity anomaly and a positive correlation between gravity and topography. The top of the lower elastic plate is bounded by the Moho, the warping of which also leads to Bouguer gravity anomalies. The light and dark shaded regions correspond to the same rheological zones represented in the upper panel.

elastic plate thickness (~30 km). However, large values for  $f$  are required to simulate folding (that is, upwarp of subsurface density discontinuities beneath topographic highs) given a model that produces isostatic compensation (downwarping of density discontinuities beneath topographic highs). The addition of up to 500 MPa (5 kbar) of compression to the plate only modifies the shape of the coherence curves slightly.

Further support for this interpretation comes from laboratory experiments on layered, analogue materials intended to simulate the rheology in Fig. 6. Burg *et al.*<sup>28</sup> report that the initial stages of compression produce periodic buckling at two wavelengths: the longer, first-order wavelength is four times the combined thickness of the two strong layers and their intermediate ductile zone, and the shorter wavelength is four times the thickness of the uppermost layer and involves deformation of that layer alone. Based on the observed scales of folding in Tibet, the predicted thickness for the upper strong layer is 35 km, and the total depth to the base of the strong region in the upper mantle is 125 km, values close to that predicted from the rheological arguments and the best-fitting elastic plate thickness (Fig. 4b). The fact that the 500-km undulations are out of phase with the surface (Fig. 5d–f) suggests that the decoupling in the lower



crust is sufficient to lead to an 'inverse boudinage' style of folding (Fig. 6).

The total amount of strain is difficult to estimate. If, based on the fit to the coherence data, the upper mantle deforms as a 50-km elastic plate, the amount of shortening obtained by 'unfolding' the 500-km wavelength undulations is at most a few per cent. In the upper-crustal layer, however, the elastic strength is likely to have been exceeded everywhere, such that it deforms as a viscous or plastic layer. The fact that the coherency at 100–200 km (Fig. 5b) involves both positive and negative correlation between gravity and topography means that folding as well as layer thickening has occurred. Thus the strain in the upper crust could be substantial, as is also the case for the lower-crustal decoupling zone.

### Testing models for Tibet

This interpretation of the Bouguer gravity and topography data for Tibet in terms of folding of layers in the upper crust and

uppermost mantle permits a reassessment of the merits of the three simple models presented in Fig. 1. Our layered strength profile with a decoupling zone in the lower crust and a rigid upper mantle is more in keeping with the predictions of the Asian subduction (Fig. 1a, c) models than with that of the rigid indenter model (Fig. 1b), in that the upper mantle appears too strong to shorten by the amount needed to double the thickness of the crust and thus must subduct. Either the Indian plate or the Asian lithosphere might be expected to display an effective elastic thickness of ~40–50 km. Thus our rigidity estimate does not distinguish the different polarities of subduction for the models shown in Fig. 1a and c but does preclude a weak upper mantle as in Fig. 1b. It is intriguing to note, however, that the arcuate gravity low centred at 34° N, 90° E (Fig. 2) is very similar to the negative gravity anomalies over oceanic trenches. If the oceanic analogue holds, the fact that the low is convex to the north would imply subduction of the Asian mantle as proposed by Willett and Beaumont<sup>7</sup>. □

Received 27 December 1993; accepted 25 August 1994.

1. Molnar, P. & Tapponnier, P. *Science* **189**, 419–426 (1975).
2. Prell, W. L. & Kutzbach, J. E. *Nature* **360**, 647–652 (1992).
3. Raymo, M. E., Ruddiman, W. F. & Froelich, P. N. *Geology* **16**, 649–653 (1988).
4. Xu, R. H., Schärer, U. & Allègre, C. J. *J. Geol.* **93**, 41–57 (1985).
5. Powell, C. M. & Conaghan, P. J. *Earth planet. Sci. Lett.* **20**, 1–12 (1973).
6. Barazangi, M. & Ni, J. *Geology* **10**, 179–185 (1982).
7. Willett, S. D. & Beaumont, C. *Nature* **369**, 642–645 (1994).
8. Dewey, J. F. & Burke, K. C. A. *J. Geol.* **81**, 683–692 (1973).
9. England, P. & McKenzie, D. P. *Geophys. J. R. astr. Soc.* **70**, 295–321 (1982).
10. Vilotte, J. P., Daignieres, M. & Madariaga, R. *J. geophys. Res.* **87**, 10709–10728 (1982).
11. England, P. C. & Houseman, G. A. *J. geophys. Res.* **94**, 17561–17579 (1989).
12. Zhao, W. & Morgan, W. J. *Tectonics* **4**, 359–369 (1985).
13. Karner, G. D. & Watts, A. B. *J. geophys. Res.* **88**, 10449–10477 (1983).
14. Lyon-Caen, H. & Molnar, P. *J. geophys. Res.* **88**, 8171–8192 (1983).
15. Sun, W. *Bouguer Gravity Anomaly Map of the People's Republic of China* (Chinese Academy of Geosurveying, Beijing, 1989).
16. Fielding, E., Isacks, B., Barazangi, M. & Duncan, C. *Geology* **22**, 163–167 (1994).

17. Bird, P. *J. geophys. Res.* **96**, 10275–10286 (1991).
18. Forsyth, D. W. *J. geophys. Res.* **90**, 12623–12632 (1985).
19. Bechtel, T. D., Forsyth, D. W., Sharpton, V. L. & Grieve, R. A. *Nature* **343**, 636–638 (1990).
20. Bechtel, T. D., Forsyth, D. W. & Swain, C. J. *Geophys. J. R. astr. Soc.* **90**, 445–465 (1987).
21. Zuber, M. T., Bechtel, T. D. & Forsyth, D. W. *J. geophys. Res.* **84**, 9353–9367 (1989).
22. Weissel, J. K., Anderson, R. N. & Geller, C. A. *Nature* **287**, 284–291 (1980).
23. McNutt, M. K., Diament, M. & Kogan, M. G. *J. geophys. Res.* **93**, 8825–8838 (1988).
24. Chen, W.-P. & Molnar, P. *J. geophys. Res.* **88**, 4183–4214 (1983).
25. Burov, E. B., Lobkovsky, L. I., Cloetingh, S. & Nikishin, A. M. *Tectonophysics* **226**, 73–87 (1993).
26. McAdoo, D. C. & Sandwell, D. T. *J. geophys. Res.* **90**, 8563–8569 (1985).
27. Zuber, M. J. *J. geophys. Res.* **92**, E541–E551 (1987).
28. Burg, J.-P., Davy, P. & Martinod, J. *Tectonics* **13**, 475–483 (1994).

ACKNOWLEDGEMENTS. We thank Y. Zhou for carefully digitizing the gravity and topography data, and L. Royden and S. Willett for insight on the models for the polarity of subduction beneath Tibet. The manuscript benefited from comments and reviews by M. Barazangi, R. Buck, D. Forsyth, P. Molnar, M. Parmentier, C. Ruppel and M. Zuber. This work was supported by NASA.

# Synergy between basic fibroblast growth factor and HIV-1 Tat protein in induction of Kaposi's sarcoma

Barbara Ensoli<sup>\*</sup>, Rita Gendelman<sup>\*</sup>, Phillip Markham<sup>†</sup>, Valeria Fiorelli<sup>\*</sup>, Sandra Colombini<sup>\*</sup>, Mark Raffeld<sup>‡</sup>, Aurelio Cafaro<sup>§</sup>, Hsiao-Kuey Chang<sup>\*</sup>, John N. Brady<sup>¶</sup> & Robert C. Gallo<sup>\*</sup>

<sup>\*</sup> Laboratory of Tumor Cell Biology, <sup>†</sup> Laboratory of Pathology and <sup>¶</sup> Laboratory of Molecular Virology, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, USA

<sup>‡</sup> Advanced BioScience Laboratories Inc., Kensington, Maryland 20895, USA

<sup>§</sup> Henry M. Jackson Foundation, Rockville, Maryland 20850, USA

**Basic fibroblast growth factor (bFGF) and human immunodeficiency virus type 1 (HIV-1) Tat protein synergize in inducing angiogenic Kaposi's sarcoma-like lesions in mice. Synergy is due to Tat, which enhances endothelial cell growth and type-IV collagenase expression in response to bFGF mimicking extracellular matrix proteins. The bFGF, extracellular Tat and Tat receptors are present in HIV-1-associated KS, which may explain the higher frequency and aggressiveness of this form compared to classical Kaposi's sarcoma where only bFGF is present.**

KAPOSI'S sarcoma (KS) is an angioproliferative disease which is rare and mild in elderly men of Mediterranean origin (classical KS) but frequent and aggressive in association with HIV-1 infection (acquired immunodeficiency syndrome (AIDS)-KS)<sup>1,3</sup>. Both forms, however, are characterized by similar histology<sup>1,3</sup>, including the presence of proliferating spindle-shaped cells considered to be the 'tumour' cells of KS, angiogenesis, inflamma-

tory cell infiltration and oedema. The spindle cells represent a heterogenous population dominated by endothelial cells mixed with cells of dendritic and monocytic origin<sup>4</sup>.

Several lines of evidence suggest that KS is a cytokine-mediated disease, at least in early stages<sup>5</sup>, and that angiogenic factors and, in particular, basic fibroblast growth factor (bFGF)<sup>6,7</sup>, play a role in lesion development. For example, spindle cells of endo-

thelial origin derived from KS lesions of AIDS patients (AIDS-KS cells)<sup>8,10</sup> induce angiogenesis in the chick chorioallantoic membrane assay and, after injection in nude mice, they induce vascular lesions of mouse cell origin closely resembling KS<sup>9</sup>. One of the prominent cytokines produced by AIDS-KS cells is bFGF, which promotes growth of these cells in an autocrine fashion<sup>10</sup> and, after release, stimulates endothelial cell migration, invasion and proliferation<sup>10,11</sup>, events required for angiogenesis<sup>6</sup>. In fact, antisense oligodeoxynucleotides directed against bFGF messenger RNA block KS cell growth and lesion formation in nude mice<sup>11</sup>. Because bFGF mRNA is expressed *in vivo* in spindle cells both of AIDS-associated and of classical KS<sup>12</sup>, these data suggest that bFGF plays a key role in the pathogenesis of both forms. However, AIDS-KS is more frequent and more aggressive than classical KS, suggesting that other factors act with bFGF in individuals infected with HIV-1. One such factor may be the HIV-1 Tat protein which is released by infected T cells<sup>13,14</sup>. Some *tat*-transgenic mice develop KS-like lesions<sup>43</sup> and extracellular Tat promotes growth, migration, invasion and adhesion of endothelial cells, precursors of KS spindle cells and AIDS-KS cells themselves<sup>13,17</sup>. However, with normal endothelial cells, the effects of Tat also require earlier cell activation with inflammatory cytokines such as  $\gamma$ -interferon, interleukin-1 and tumour necrosis factor (refs 15–17 and V.F. *et al.*, manuscript submitted). These cytokines induce endothelial cells to acquire features of KS cells, including the spindle morphology and the expression of the integrin receptors  $\alpha_5\beta_1$  and  $\alpha_v\beta_3$  which, in turn, mediate the effects of extracellular Tat on these cells<sup>15,16</sup>. The same cytokines are increased in individuals infected with HIV-1<sup>15,18,19</sup>, and homosexual men, who are at highest risk for developing KS, show evidence of immune activation and may have higher cytokine levels even before HIV-1 infection<sup>15,20</sup>. Inflammatory cytokines also activate production and release of bFGF from endothelial cells<sup>21,22</sup>, suggesting a mechanism for the increased bFGF expression in KS.

Here we provide evidence that bFGF and Tat synergize in inducing KS development and that extracellular Tat may represent the factor increasing the frequency and aggressiveness of KS in individuals infected with HIV-1.

### bFGF, Tat and integrins in AIDS-KS lesions

To verify whether bFGF and Tat are present in KS, frozen sections from AIDS-associated KS lesions, classical KS lesions and control tissues were stained with affinity-purified antibodies directed against these proteins. Factor VIII-related antigen (FVIII-RA) was monitored to identify vessels and endothelial spindle cells. HIV-1 p24 core antigen was evaluated to identify virus infected cells. bFGF was detected in all AIDS-KS (10/10) and classical KS lesions (2/2) examined in both vessels and spindle cells (Table 1a), whereas bFGF was present only around the vessels in control tissues probably stored in the basement membrane<sup>23</sup> (Fig. 1a). Anti-FVIII-RA antibodies stained the vessels in all tissues (KS and controls) and many of the spindle cells in both AIDS-KS and classical KS (Fig. 1b). The endothelial origin of these spindle cells was also confirmed by using antibodies directed against cadherin-5, CD34, EN-4 and ELAM-1 (data not shown). These results indicated that bFGF is expressed by spindle cells of both AIDS-associated and classical KS and suggested that most of the bFGF-producing spindle cells are endothelial.

The same tissues were then stained with anti-Tat and anti-p24 antibodies. Tat was detected in 7/10 and p24 in 4/7 AIDS-KS lesions examined, but not in classical KS or control tissues (Table 1a). Stronger Tat staining was associated with more advanced lesions and was localized more often in spindle cells whereas p24 was mostly positive in mononuclear cells (Fig. 1c, d). This confirmed previous data from *in situ* hybridization on the presence of cells infected with HIV-1 in AIDS-KS<sup>24</sup> and raised the question of the nature of the Tat-positive spindle cells. In particular, these cells may be uninfected cells that have taken

up Tat released by infected cells, as found *in vitro* with KS spindle cells and normal endothelial cells<sup>14</sup> or in mice injected with Tat<sup>25</sup>. To verify this, an AIDS-KS lesion was doubly stained with anti-Tat and anti-FVIII-RA, -CD14, -CD4, -p24 or -HIV-1 reverse transcriptase (RT) antibodies combined. Over 50% of the total Tat-positive cells (22%) were uninfected endothelial spindle cells (Table 1b). Spindle-shaped and mononuclear HIV-1-infected macrophages and infected lymphocytes represented 10%, 5% and 33% of the total Tat-positive cells, respectively. Other HIV-1-infected mononuclear cells (9%) (CD14+, p24+ and RT+) did not express detectable Tat (data not shown). Thus, Tat released by infected cells<sup>13,14</sup> is taken up by KS spindle cells *in vivo*, as found *in vitro*<sup>14</sup>.

Uptake of Tat by KS cells or cytokine-activated endothelial cells appears to be mediated by the same integrin receptors (A. Holmes *et al.*, manuscript in preparation) that are highly expressed by KS spindle cells and are inducible by inflammatory cytokines on endothelial cells<sup>16</sup>. These integrins,  $\alpha_5\beta_1$  and  $\alpha_v\beta_3$ , function as the receptors for Tat and mediate its effects on these cell types by binding to the RGD (arginine-glycine-aspartic acid) sequence of the protein<sup>16</sup>, as they normally do with extracellular matrix (ECM) proteins, such as fibronectin (FN) and vitronectin (VN)<sup>26</sup>. The same integrins are expressed by endothelial cells *in vivo* during angiogenesis induced by bFGF or inflammatory cytokines<sup>27</sup>. Therefore, KS lesions and control tissues were stained with antibodies recognizing the RGD-binding region of the  $\alpha_5\beta_1$  and  $\alpha_v\beta_3$  receptors<sup>16</sup> alone or combined with anti-Tat antibodies. Both integrins were found to be highly expressed by spindle cells and vessels of AIDS-KS (Table 1c). Classical KS showed lower positivity with prominent  $\alpha_v\beta_3$  staining, whereas in control tissues only some vessels were positive. In addition, Tat-positive spindle cells of AIDS-KS also stained for  $\beta_1$  or  $\beta_3$  integrins confirming the presence of both Tat and its receptors on the same cells.

### bFGF induces angiogenic KS-like lesions

To determine whether bFGF can promote the histological changes found in KS, increasing amounts of the protein were injected subcutaneously into nude mice and compared to the effect of AIDS-KS cells (Table 2 and Fig. 2a). Human umbilical vein endothelial (H-UVE) cells, bovine serum albumin (BSA), and buffer or media were used as negative controls (Table 2). bFGF, but not BSA or buffer, induced the formation of macroscopic vascular lesions at the site of injection in a dose-dependent manner (Fig. 2a and Table 2). Specifically, 0.1  $\mu$ g bFGF induced only microscopic alterations, 1  $\mu$ g bFGF induced macroscopic lesions in 44% of the animals and 10  $\mu$ g or greater amounts induced macroscopic lesions in 100% of the mice. Lesions were very similar to those induced by AIDS-KS cells, both developing within 3–4 days and reaching their maximum size by day 6–7. Histologically, both lesions presented alterations characteristic of early stage KS such as angiogenesis, spindle cell proliferation, oedema and inflammatory cell infiltration (Table 2 and Fig. 2A). The frequency and the intensity of these features depended on the amount of bFGF inoculated (Table 2), except for acute (neutrophilic) and chronic (mononuclear) inflammatory cell infiltration and oedema which were always less prevalent with bFGF compared to AIDS-KS cells (Table 2). This is consistent with data indicating that AIDS-KS cells also produce cytokines mediating inflammatory effects<sup>10</sup>.

### bFGF and Tat synergize to induce KS-like lesions

Because bFGF is expressed at high levels both in AIDS-associated and classical KS<sup>10,12</sup> (Table 1 and Fig. 1), it is likely that this cytokine contributes to the histogenesis of both forms. However, AIDS-KS is more frequent and aggressive than classical KS. As Tat and bFGF are both present in AIDS-KS (Table 1 and Fig. 1), the role of Tat in KS lesion formation was investigated by injecting nude mice with Tat protein alone or combined with bFGF (Table 3). In the absence of bFGF, Tat stimulated histo-



logical alterations similar to those found with 0.1  $\mu\text{g}$  bFGF (Fig. 2*Ba, b*). These alterations were evident with 1–10  $\mu\text{g}$  Tat and did not increase with higher doses (up to 100  $\mu\text{g}$ ; data not shown). This is consistent with the requirement of other factors (that is, inflammatory cytokines) inducing the receptors for Tat<sup>14–17</sup> and with recent findings indicating that the effects of

Tat on KS and endothelial cells may be mediated by mechanism(s) other than the direct transactivation of gene expression, which is dose-dependent and requires much higher concentrations of the protein<sup>13–17,28</sup>. However, when Tat and bFGF were injected simultaneously, both the incidence of macroscopic lesions and the histopathology markedly increased as compared

TABLE 1 Staining of bFGF, FVIII-RA, HIV-1 Tat and p24 (a), characterization of Tat-positive cells (b) and  $\alpha_5\beta_1$ ,  $\alpha_v\beta_3$  integrin (Tat) receptor expression (c) in AIDS-KS, classical KS and control tissues

(a) Staining results

| Specimen     | bFGF*      | Average % (range) positive cells |            | p24        |
|--------------|------------|----------------------------------|------------|------------|
|              |            | FVIII*                           | TAT        |            |
| AIDS-KS      |            |                                  |            |            |
| Lymph node   | 28 (20–34) | 34 (21–45)                       | 22 (13–31) | 24 (11–33) |
| Skin         |            |                                  |            |            |
| 1            | 19 (17–20) | 39 (35–44)                       | 27 (12–38) | 23 (21–25) |
| 2            | 22 (16–30) | 47 (30–78)                       | 18 (12–32) | 13 (10–15) |
| 3            | 33 (21–42) | 49 (27–67)                       | 12 (6–18)  | 16 (11–21) |
| 4            | 27 (23–31) | 41 (36–47)                       | —          | —          |
| 5            | 4 (2–6)    | 10 (6–13)                        | —          | —          |
| 6            | 15 (10–28) | 68 (59–76)                       | 7 (3–10)   | ND         |
| 7            | 29 (21–43) | 51 (41–58)                       | 12 (10–13) | ND         |
| 8            | 7 (5–9)    | 39 (34–44)                       | 8 (5–11)   | ND         |
| 9            | 10 (6–14)  | 18 (14–22)                       | —          | —          |
| Classical KS |            |                                  |            |            |
| Skin         |            |                                  |            |            |
| 1            | 28 (26–30) | 41 (33–50)                       | —          | —          |
| 2            | 9 (7–10)   | 30 (24–34)                       | —          | —          |
| Controls     |            |                                  |            |            |
| Tonsil       |            |                                  |            |            |
| 1            | —          | —                                | —          | —          |
| 2            | —          | —                                | —          | —          |
| Skin         |            |                                  |            |            |
| 1            | —          | —                                | —          | —          |
| 2            | —          | —                                | —          | —          |
| 3            | —          | —                                | —          | —          |

(b) AIDS-KS lesion (lymph node)

| Tat-Positive Cells<br>% (Range) | Phenotype<br>FVIII-RA | CD14 | CD4 | p24 | RT | Morphology  | HIV-1 infection |
|---------------------------------|-----------------------|------|-----|-----|----|-------------|-----------------|
| 11% (7–16)                      | +                     | —    | —   | —   | —  | Spindle     | —               |
| 2% (0.5–3)                      | —                     | +    | +   | +   | +  | Spindle     | +               |
| 7% (5–9)                        | —                     | —    | +   | +   | +  | Mononuclear | +               |
| 1% (0–2)                        | —                     | +    | +   | +   | +  | Mononuclear | +               |

(c) Average % (range) of  $\alpha_5\beta_1$  and  $\alpha_v\beta_3$  positive spindle cells and double positive for Tat†

| Specimen†    | $\alpha_5$ | $\beta_1$  | $\alpha_v$ | $\beta_3$  | $\beta_1/\text{Tat}$ | $\beta_3/\text{Tat}$ | Tat       |
|--------------|------------|------------|------------|------------|----------------------|----------------------|-----------|
| AIDS-KS      | 24 (13–35) | 15 (10–34) | 41 (21–52) | 50 (34–69) | 5 (3–7)              | 12 (10–13)           | 14 (9–19) |
| Classical KS | 3 (1–7)    | 3 (1–13)   | 24 (15–32) | 25 (13–48) | —                    | —                    | —         |
| Tonsil       | —          | —          | —          | —          | —                    | —                    | —         |
| Control skin | —          | —          | —          | —          | —                    | —                    | —         |

Frozen sections from skin and tonsil of individuals not infected with HIV-1, from involved skin and lymph node of AIDS-KS patients and from involved and uninvolved skin of classical KS patients were fixed in cold acetone and single- or double-stained by the alkaline-phosphatase-anti-alkaline phosphatase (APAAP) and/or the peroxidase-anti-peroxidase (PAP) methods<sup>14,16</sup> by using affinity-purified mouse monoclonal or rabbit polyclonal antibodies alone or combined (APAAP revealed the monoclonal and PAP the polyclonal antibodies). For single-staining with APAAP, slides were incubated (20 min, room temperature) with rabbit anti-mouse antibodies (1:25) (Dako, Glostrup, Denmark). After washing, the APAAP complex (1:25) was applied for 20 min. The second and third steps were repeated twice and the conjugate was preadsorbed with standard immunoglobulins (Boehringer-Mannheim Standards). For PAP and double-stainings (APAAP/PAP) the universal Dako double-stain kit system 40 was used by manufacturer protocol. Monoclonal antibodies were: anti-bFGF (1:200) (Promega); anti-FVIII-RA (1:100) (Dako); anti-Tat<sup>14</sup> (1:250); anti-Tat (1:1,000) and anti-RT (1:50) (American Biotechnology); anti-p24 (1:5), anti-CD14 (1:20) and anti-CD4 (1:10) (Dako); anti- $\alpha_5$  (CDW49e), anti- $\beta_1$  (CD29), anti- $\alpha_v$  (CD51) and anti- $\beta_3$  (CD61) (1:50) (AMAC, Cambridge, MA). Polyclonal antibodies were: anti-FVIII (1:500) (Dako), anti-CD4 (1:250) and anti-p24 (1:250) (American Biotechnology) and anti-Tat<sup>14</sup> (1:500). All antibodies were characterized for specificity before use. Stainings were done with each antibody directed against the same antigen with identical results. All combinations of monoclonal and polyclonal antibodies were used for double stainings. Reported are the average results and the range (minimal and maximal values) of the percent of positive cells in 3–10 high-power microscopic fields from 3 or more experiments. Results were also confirmed by immunofluorescence.

\* Vessels stained positive for FVIII and bFGF in all tissues.

† Specimens were 2 AIDS-KS (lymph node and skin), 4 classical KS (skin), 1 tonsil, 2 normal skin.

‡ All positive cells had spindle morphology. Most vessels were positive for both integrins with prevalence for  $\alpha_v\beta_3$  in all KS tissues analysed with highest positivity in AIDS-KS; some vessel positivity was seen in control tissues.

ND, Not done; (—), negative; (+), positive.

with each protein alone (Table 3 and Fig. 2*Bc*). For example, when Tat (10 µg) and bFGF (0.1 or 1 µg) were combined, macroscopic lesions developed in 35% and 71% of the mice, respectively, reaching effects similar to 1 µg and 10 µg bFGF alone (Table 3). Tat also increased the intensity of the histological alterations (particularly spindle cell growth, angiogenesis

and oedema) for each dose of bFGF inoculated, reaching values generally observed with about 10-fold higher bFGF concentrations (Table 3 and Fig. 2*B*). In contrast, heat-inactivated Tat had no synergistic effects with bFGF (0.1 or 1 µg) (Table 3). Similarly, synergy was blocked by anti-bFGF or anti-Tat antibodies (Table 3). Because after each injection in mice Tat was

TABLE 2 Injection of bFGF in nude mice induces KS-like lesions in a dose-dependent fashion

| bFGF                              | Lesion*   | Angiogenesis† | Spindle cells† | Acute inflammation† | Chronic inflammation† | Oedema†  |
|-----------------------------------|-----------|---------------|----------------|---------------------|-----------------------|----------|
| 0.1 µg                            | 0% (47)   | 32% (1)       | 75% (1)        | 50% (1)             | 60% (1)               | 75% (1)  |
| 1 µg                              | 44% (43)  | 86% (2.7)     | 95% (3)        | 62% (1)             | 62% (2)               | 76% (2)  |
| 10 µg                             | 100% (11) | 100% (5)      | 100% (5)       | 70% (1)             | 100% (2)              | 100% (2) |
| 30 µg                             | 100% (19) | 100% (6)      | 100% (6)       | 100% (1)            | 100% (2)              | 100% (2) |
| 90 µg                             | 100% (10) | 100% (7)      | 100% (7)       | 100% (1)            | 100% (2)              | 100% (2) |
| BSA 90 µg                         | 0% (2)    | 0%            | 0%             | 0%                  | 0%                    | 0%       |
| Buffer                            | 0% (36)   | 0%            | 0%             | 0%                  | 0%                    | 0%       |
| AIDS-KS cells 3 × 10 <sup>6</sup> | 100% (10) | 100% (4)      | 100% (4)       | 100% (6)            | 100% (4)              | 100% (3) |
| H-UVE cells 3 × 10 <sup>6</sup>   | 0% (6)    | 0%            | 0%             | 0%                  | 0%                    | 0%       |

Recombinant purified human bFGF (Boehringer-Mannheim), BSA (Sigma), AIDS-KS cells or H-UVE cells ( $3 \times 10^6$ )<sup>8-10</sup> were injected subcutaneously into the lower back (right side) of BALB/c nu/nu athymic mice<sup>9</sup>, whereas their negative controls (protein buffer [PBS, 0.1% BSA] or media) were injected into the left side of the same mice. To increase efficiency, proteins, cells or buffer (200 µl) were mixed with 200 µl of matrigel (Collaborative Research, Bedford, MA)<sup>41</sup> before inoculation. Mice were killed at 6–7 days after injection. With bFGF or AIDS-KS cells, lesions developed within 3–4 days reaching maximal size ( $4 \times 5$  to  $9 \times 12$  mm) by day 6–7, whereas none of the 86 animals inoculated with the negative controls (buffer or media in matrigel) developed lesions. Tissue samples were fixed in formalin and analysed microscopically after a haematoxylin-eosin (H and E) staining. The histological changes observed at the site of injection, blood vessel formation, spindle cell proliferation, acute (neutrophilic) and chronic (mononuclear) inflammatory cell infiltration and oedema, were evaluated by comparison with the negative controls and graded according to intensity with values ranging from 1–7 with the minimal alteration observed given a value of one (intensity value). Each intensity value reported is the average result from all mice presenting that alteration per experimental condition. Vascular lesions and histological changes were also observed in the absence of matrigel.

\* Percentage of mice developing vascular macroscopic lesions. The number of mice inoculated is in parentheses.

† Percentage of mice inoculated developing histological alterations. Average 'intensity' value for each histopathological feature observed per experimental condition is in parentheses.

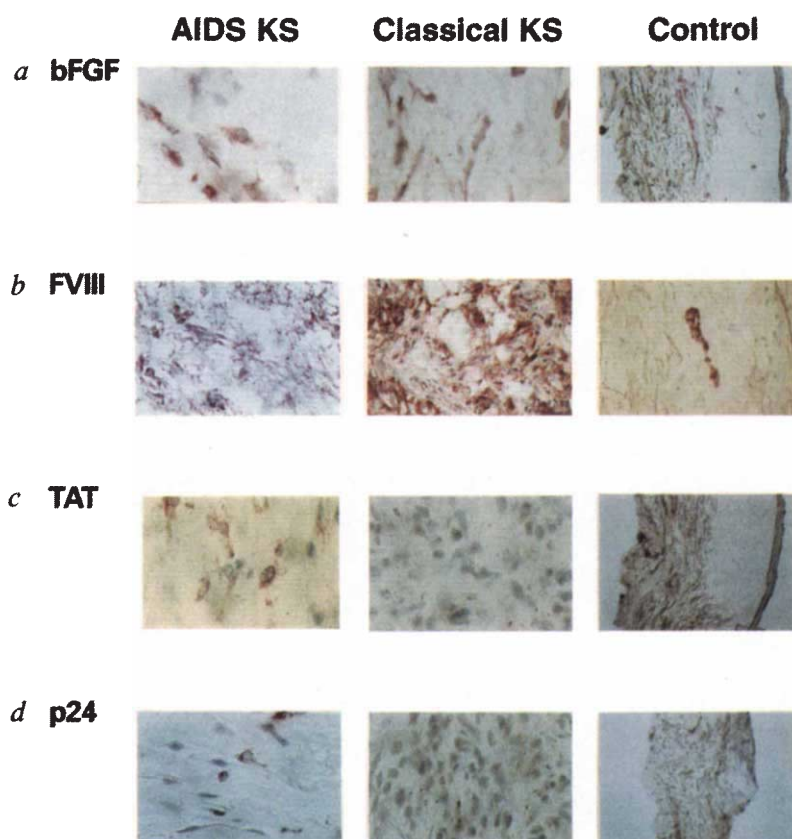
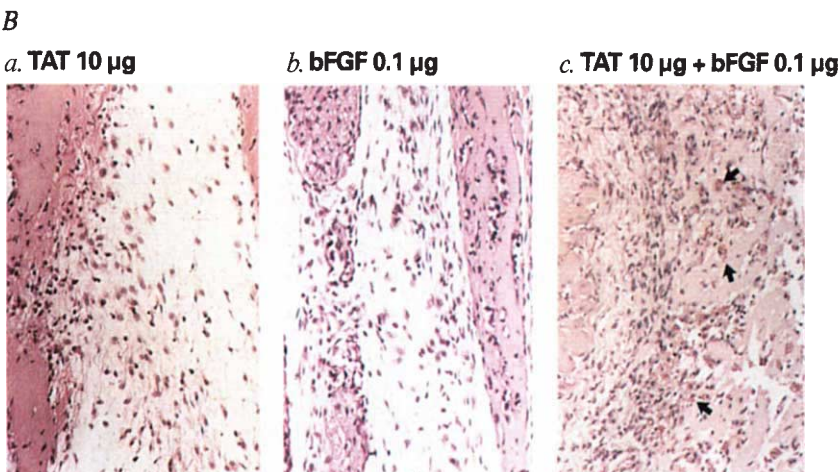
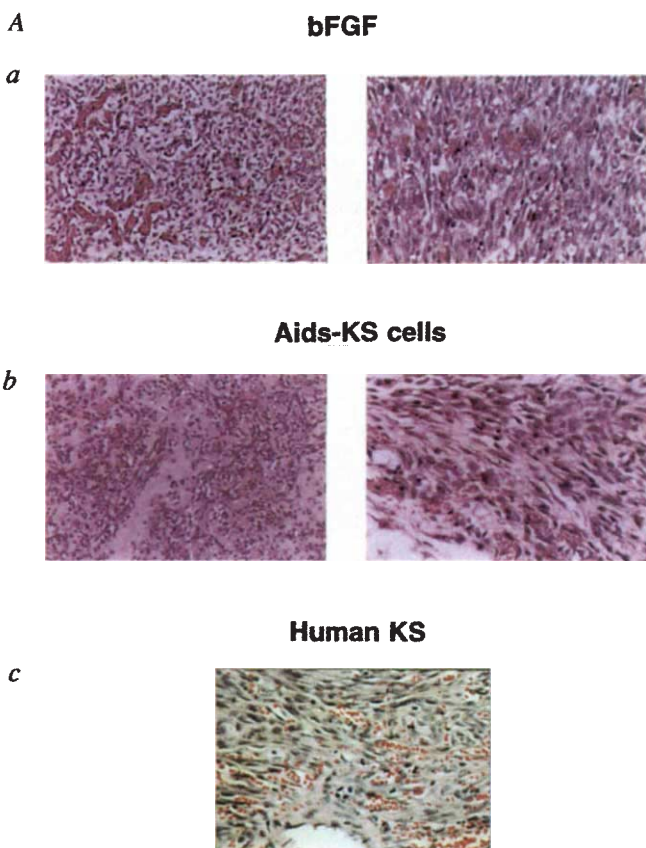


FIG. 1 bFGF, FVIII-RA, HIV-1 Tat and HIV-1 p24 staining in AIDS-KS, classical KS and control tissues. *a*, bFGF, *b*, FVIII-RA, *c*, TAT and *d*, p24 staining of a representative AIDS-KS skin lesion (left panels), classical KS skin lesion (middle panels) and control skin (right panels), respectively. Shown are examples of the staining with the antibodies reported in Table 1*a*. Magnifications are  $\times 22$  or  $\times 54$  for KS lesions and  $\times 5.4$  or  $\times 22$  for control skin. For methodology see the legend to Table 1.



fully active, these data indicate that biologically active protein is needed for the synergistic KS-promoting effect with bFGF and that the effects of both Tat and bFGF are specific.

To investigate the mechanism of this synergistic effect, bFGF, Tat or buffer were injected on day 0 and the other protein of the combination was inoculated in the same sites 2 days later. When bFGF was injected before Tat, synergy was more evident than the simultaneous inoculation, reaching 45% and 100% of lesion formation with 0.1 or 1  $\mu\text{g}$  bFGF and Tat, respectively. In contrast, when Tat was injected on day 0 and bFGF on day 2, no synergy occurred (supplementary information available non request). This indicates that bFGF is essential for lesion formation and that Tat is the factor enhancing bFGF activity. This observation may partially explain the high frequency of KS in homosexual men who may have increased levels of inflammatory cytokines activating bFGF expression<sup>21,22</sup> even before HIV-1 infection<sup>15,18-20</sup>.



## Tat mimics the effect of ECM molecules

There may be several mechanisms by which Tat enhances bFGF activity. Endothelial cells require two signals in order to grow during angiogenesis: one from angiogenic factors, such as bFGF, which trigger growth and integrin expression, and one from ECM molecules, such as FN, which induce cell adhesion and modulate cell responsiveness to bFGF<sup>27,29</sup>. FN regulates endothelial cell growth (both basal and bFGF-induced) by modulating cell shape and by activating intracellular signalling pathways mediated through its RGD sequence and the  $\alpha_5\beta_1$  receptor<sup>29</sup>. Our previous data indicated that, like ECM molecules, Tat mediates vascular and AIDS-KS cell adhesion by the interaction of its RGD sequence with the integrins  $\alpha_5\beta_1$  and  $\alpha_v\beta_3$ , the receptors for FN and VN, respectively<sup>16</sup>, and that this effect is increased by the basic region of the protein which interacts with heparin complexes of the cell surface (ref. 16 and our unpublished data).

To determine whether Tat could mimic the effect of ECM molecules on endothelial cell growth, H-UVE cells were seeded onto non-adhesive plates coated with Tat (0.01, 0.1 or 1  $\mu\text{g ml}^{-1}$ ), 0.1% BSA (negative control) or 30  $\mu\text{g ml}^{-1}$  FN (positive control). The day after, the percentage of adherent cells was counted, and then [<sup>3</sup>H]thymidine incorporation was evaluated 48 h after the addition of bFGF (0.01, 0.1 or 1  $\text{ng ml}^{-1}$ ) or buffer. Tat could support cell adhesion and spreading in a dose-dependent fashion and at 1  $\mu\text{g ml}^{-1}$  reached the effect of equimolar concentrations of FN (Fig. 3Aa). Tat also increased both basal and bFGF-induced cell growth at the same levels observed with FN (Fig. 3Ab, c). Thus, Tat can support endothelial cell adhesion, spread and growth, and can increase the cell growth response to bFGF mimicking the effect of FN<sup>29</sup>. As Tat receptors are expressed by endothelial and spindle cells of KS (Table 1c) and Tat can compete against endothelial and KS spindle cell adhesion to FN and VN<sup>16</sup>, these results suggest that Tat can exert these activities *in vivo*.

## FN and TAT have similar effects *in vivo*

To determine whether FN could induce effects similar to Tat *in vivo*, the molecule (30  $\mu\text{g}$ ) was injected into mice alone or with bFGF (0.1 or 1  $\mu\text{g}$ ). As found with Tat, FN alone did not induce lesions in mice; however, in the presence of bFGF, lesions developed in 40% and 80% of the mice, respectively (supplementary information available on request). When bFGF was injected on day 0 and FN on day 2, lesion development was further enhanced to 60% and 100% of the mice, respectively. These results indicated that Tat can mimic the effect of ECM proteins such as FN both *in vitro* and *in vivo*.

FIG. 2 A, Injection of bFGF into nude mice induces vascular lesions closely resembling the lesions induced by AIDS-KS spindle cells or primary human KS. Different fields ( $\times 200$  and  $\times 400$  magnifications) of the histology of the mouse lesions induced by bFGF (a) or AIDS-KS cells (b), compared to a human primary KS lesion (c) ( $\times 400$  magnification). Slides were stained by H and E. See the legend to Table 2 for methodology. Angiogenesis and spindle cells are evident in all panels, whereas oedema and inflammatory cell infiltration are more pronounced after injection of AIDS-KS cells (b) and in human KS (c). B, Tat and bFGF have synergistic effects in inducing KS-like lesions. H&E staining of representative tissues samples from mice injected with a, 10  $\mu\text{g}$  Tat; b, 0.1  $\mu\text{g}$  bFGF; c, 10  $\mu\text{g}$  Tat and 0.1  $\mu\text{g}$  bFGF. Magnifications,  $\times 86$ . See the legends to Tables 2 and 3 for experimental procedures. Arrows point to angiogenesis. In the presence of Tat and bFGF combined, both the frequency and the intensity of the histological alterations (particularly angiogenesis, spindle cell growth and oedema) are increased as compared to each protein alone.

## Activation of type-IV collagenase expression

Picomolar concentrations ( $10\text{--}20\text{ ng ml}^{-1}$ ) of Tat induce endothelial cells to migrate and to invade the basement membrane and this is associated with degradation of collagen IV<sup>17</sup>. The degradation of the basement membrane can be mediated by several metalloproteinases<sup>30,31</sup>. Among them we investigated type-IV collagenase mRNA encoding the 72K species because this form is involved in angiogenesis and tumour invasion<sup>31</sup>, and is induced by bFGF<sup>32</sup>. Tat enhanced collagenase mRNA expression in endothelial cells which peaked at  $10\text{ ng ml}^{-1}$  Tat (2-fold increase; Fig. 3*Ba*). Phorbol-myristyl-acetate (PMA) (positive control)<sup>33</sup> resulted in a 2.8-fold activation (Fig. 3*B*). The effect of Tat ( $10\text{ ng ml}^{-1}$ ) was increased synergistically by the presence of 0.1 or 1 ng bFGF (Fig. 3*Bb*), and reached 2.6- and 7.2-fold increase compared to bFGF alone which induced 1.2- and 3-fold activation, respectively (Fig. 3*Bb*). Thus, bFGF and Tat synergize in inducing collagenase IV gene expression in endothelial cells.

The most likely mechanism for Tat-activation of type-IV collagenase is the binding to the  $\alpha_5\beta_1$  and  $\alpha_v\beta_3$  integrins<sup>16</sup> which activate this enzyme<sup>34-36</sup>. In contrast, direct transactivation of gene expression by Tat requires higher concentrations of extracellular protein<sup>14,28</sup> than those needed for stimulating collagenase expression and endothelial cell invasion. Because collagenase IV is involved in the degradation of the ECM and because the effect of Tat is synergistic with bFGF, the data suggest another pathway by which bFGF and Tat may have synergistic KS promoting effects *in vivo*.

The data presented here indicate that both Tat and bFGF are present in AIDS-KS lesions and that they synergize in inducing angiogenesis and histological alterations typical of KS<sup>1-5</sup>. As Tat is released from infected cells<sup>13,14</sup> and its receptors (integrins  $\alpha_5\beta_1$  and  $\alpha_v\beta_3$ )<sup>16</sup> are expressed by spindle cells and endothelial

cells of KS lesions, Tat may be the factor responsible for the higher frequency and aggressiveness of KS in individuals infected with HIV-1 as compared to the classical form of KS in which only bFGF is present.

Tat synergizes with bFGF by at least two pathways: increased endothelial cell growth and invasion in response to bFGF, mimicking the effect of ECM proteins such as FN and VN.

A possible explanation for the lack of *in vivo* effect of Tat or FN injected alone in mice is that their receptors are already occupied by endogenous ECM proteins. However, simultaneous injection of bFGF triggers angiogenesis and integrin expression<sup>27</sup>, increasing the availability of the receptors for these proteins which now are able to bind and to increase bFGF activity in a synergistic fashion. In fact, injection of bFGF before Tat or FN further increases lesion development. As bFGF and Tat are both present in AIDS-KS lesions and integrins are over-expressed in both vessels and spindle cells, it is likely that synergy by Tat occurs in individuals infected with HIV-1.

It is still unclear why KS is more frequent in infected homosexual males compared to other infected people. This could be due to greater immune activation and consequent inflammatory cytokine production in these individuals<sup>18,20</sup>. Inflammatory cytokines induce endothelial cells to acquire the features of the KS cell phenotype, including morphology<sup>15</sup>, bFGF production and release<sup>21,22</sup>, induction of Tat receptors and responsiveness to its angiogenic effects<sup>15-17</sup>. Furthermore, cytokine-activated endothelial and KS spindle cells express adhesion molecules which mediate contact with T cells, macrophages, and dermal dendrocytes<sup>4,37,38</sup>, all of which are targets of HIV-1 infection<sup>24,39,40</sup>. Because infected cells release Tat and in AIDS-KS lesions they are in contact with spindle cells, extracellular Tat may cooperate with bFGF in increasing KS lesion formation and progression. □

TABLE 3 bFGF and the HIV-1 Tat protein have synergistic angiogenic and KS-like promoting effects

|                                                                      | Lesion*   | Angiogenesis† | Spindle cells† | Acute inflammation† | Chronic inflammation† | Oedema†   |
|----------------------------------------------------------------------|-----------|---------------|----------------|---------------------|-----------------------|-----------|
| Tat 10 $\mu\text{g}$ + buffer                                        | 0% (17)   | 41% (1)       | 94% (1)        | 57% (1)             | 100% (1)              | 94% (1)   |
| bFGF 0.1 $\mu\text{g}$ + buffer                                      | 0% (47)   | 32% (1)       | 75% (1)        | 50% (1)             | 60% (1)               | 75% (1)   |
| Tat 10 $\mu\text{g}$ + bFGF 0.1 $\mu\text{g}$                        | 35% (31)  | 58% (3)       | 84% (2)        | 80% (1)             | 93% (1)               | 81% (2.5) |
| bFGF 0.1 $\mu\text{g}$ + heat-inactivated Tat 10 $\mu\text{g}$       | 0% (5)    | 60% (1)       | 60% (2)        | 60% (1)             | 60% (1)               | 60% (1)   |
| bFGF 0.1 $\mu\text{g}$ + Tat 10 $\mu\text{g}$ + anti-Tat antibodies  | 0% (10)   | 60% (1)       | 80% (1)        | 80% (1)             | 80% (1)               | 80% (1)   |
| bFGF 0.1 $\mu\text{g}$ + anti-bFGF antibodies + Tat 10 $\mu\text{g}$ | 0% (5)    | 20% (1)       | 40% (1)        | 40% (1)             | 40% (1)               | 60% (1)   |
| bFGF 1 $\mu\text{g}$ + buffer                                        | 44% (43)  | 86% (2.7)     | 95% (3)        | 62% (1)             | 62% (2)               | 76% (2)   |
| Tat 10 $\mu\text{g}$ + bFGF 1 $\mu\text{g}$                          | 71% (31)  | 94% (3.5)     | 100% (4)       | 100% (1)            | 100% (2)              | 100% (3)  |
| bFGF 1 $\mu\text{g}$ + heat-inactivated Tat 10 $\mu\text{g}$         | 40% (5)   | 80% (2)       | 80% (3)        | 80% (1)             | 80% (2)               | 80% (2)   |
| bFGF 1 $\mu\text{g}$ + Tat 10 $\mu\text{g}$ + anti-Tat antibodies    | 40% (5)   | 40% (2)       | 80% (3)        | 60% (1)             | 60% (2)               | 80% (2)   |
| bFGF 10 $\mu\text{g}$ + buffer                                       | 100% (11) | 100% (5)      | 100% (5)       | 70% (1)             | 100% (2)              | 100% (2)  |
| Tat 10 $\mu\text{g}$ + bFGF 10 $\mu\text{g}$                         | 100% (2)  | 100% (6)      | 100% (6)       | 100% (1)            | 100% (3)              | 100% (5)  |
| Buffer + buffer                                                      | 0% (70)   | 0%            | 0%             | 0%                  | 0%                    | 0%        |

Tat (10  $\mu\text{g}$ ), heat-inactivated Tat (10  $\mu\text{g}$ ), increasing amounts of bFGF (0.1, 1, 10  $\mu\text{g}$ ), Tat/anti-Tat antibodies, bFGF/anti-bFGF antibodies or protein buffer (with or without antibodies) were injected alone or in combination into mice as described in the legend to Table 2. METHODS: recombinant purified Tat protein, expressed in *E. coli*, was resuspended in the same buffer used for bFGF, handled and tested for activity as previously described<sup>14</sup>. As Tat oxidizes easily and loses biological activity<sup>13-17</sup>, the protein was tested<sup>14</sup> before and after each injection to insure that fully active Tat was inoculated. Tat protein was inactivated by incubation at  $56^\circ\text{C}$  for 5 h. Biological activity was 5–7% of the initial activity. To block Tat activity, affinity-purified anti-Tat polyclonal antibodies<sup>13,14</sup> (300  $\mu\text{l}$  at  $1.7\text{ }\mu\text{g }\mu\text{l}^{-1}$ ) were incubated with 10  $\mu\text{g}$  of Tat (at  $4^\circ\text{C}$ , in the dark) for 6–12 h with gentle shaking. Antibodies incubated in protein buffer and Tat alone were used as negative and positive controls, respectively. After incubation, bFGF and matrigel were added to the tubes and the mixtures injected in mice. *In vitro* testing<sup>14</sup> indicated that more than 96% of Tat activity was blocked by the antibodies. To block bFGF, 1 mg neutralizing affinity-purified anti-bFGF polyclonal antibodies (R&D Systems, Minneapolis, MN) were incubated with 0.1  $\mu\text{g}$  bFGF in 300  $\mu\text{l}$  total volume, or with buffer. Then Tat and matrigel were added and mixtures inoculated in mice. A second injection of the antibody (1 mg) was done in the same sites at day 2. Antibodies had no effects when injected alone (data not shown).

\* Percentage of mice developing macroscopic vascular lesions. In parentheses, number of mice inoculated.

† Percentage of mice inoculated developing histological alterations. In parentheses, average 'intensity' value for each histopathological feature observed per experimental condition.



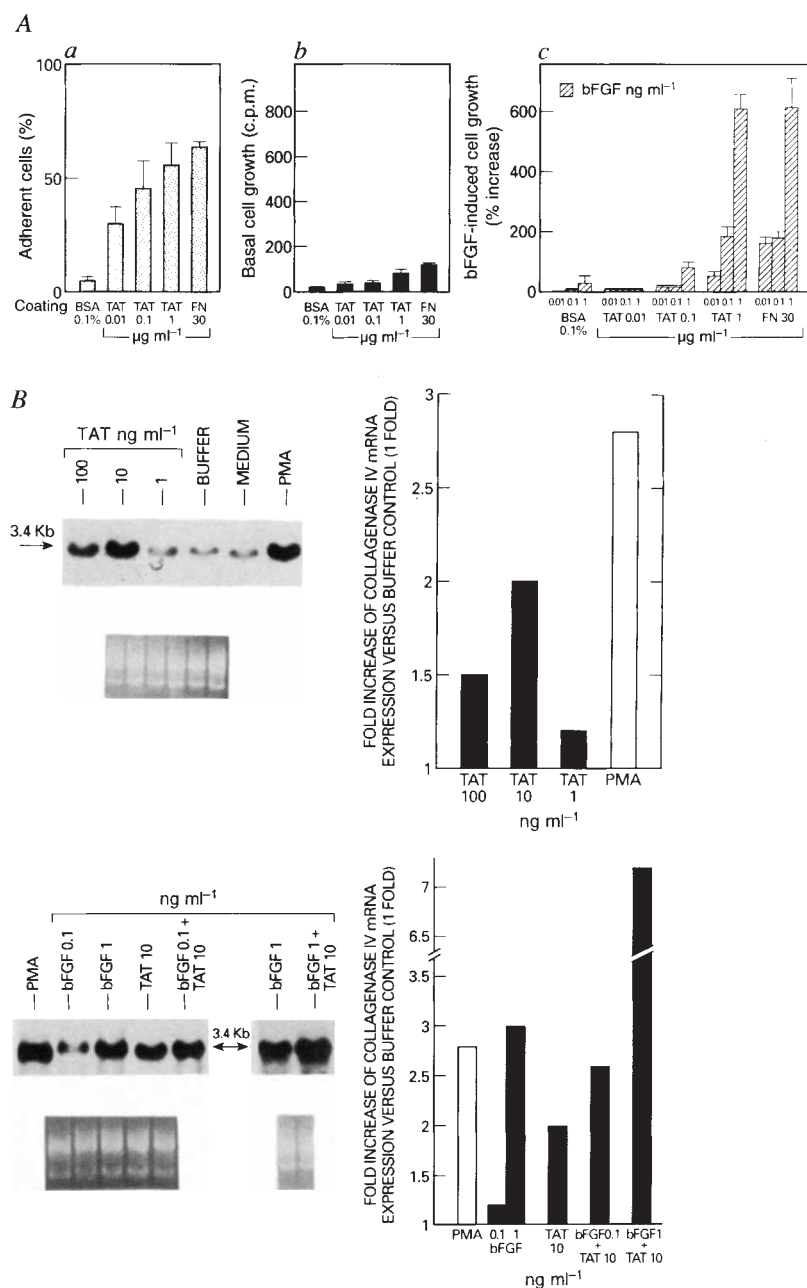


FIG. 3 A, Tat supports endothelial cell adhesion and increases growth to bFGF mimicking the effect of FN. a, Percentage adhesion; b, basal growth (c.p.m.); and c, bFGF-induced growth (percent increase) of H-UVE cells seeded onto BSA-, Tat- or FN-coated plates (average of 3 experiments). B, HIV-1 Tat protein activates endothelial type-IV collagenase expression. b, Synergistic collagenase IV activation by bFGF and Tat. Shown are Northern blots (left panels) and fold-increase of collagenase IV mRNA (over control buffer) by densitometry (right panels) with H-UVE cells treated with Tat (100, 10 or 1  $\text{ng ml}^{-1}$ ), buffer, medium or PMA (20  $\text{ng ml}^{-1}$ ) (a); bFGF (0.1 and 1  $\text{ng ml}^{-1}$ ) and Tat (10  $\text{ng ml}^{-1}$ ), alone or combined, and PMA (b). Ethidium-stained gels are shown to confirm equivalent RNA amounts.

**METHODS.** H-UVE cells (p4-p9) (Cell System Inc., Kirkland, WA) were cultured as described previously<sup>8,10,13-17</sup>. Non-tissue culture-treated 96-well plates (Linbro, Flow Laboratories, McLean, VA) were coated with 100  $\mu\text{l}$  PBS-BSA 0.1% with or without Tat (0.01, 0.1 or 1  $\mu\text{g ml}^{-1}$ ) or FN (30  $\mu\text{g ml}^{-1}$ ) in the dark (4 °C). After 16–20 h, plates were washed, 200  $\mu\text{l}$  PBS-BSA 1% added for 1 h (4 °C) and 30 min (room temperature), and washed again. Cells ( $8 \times 10^2$ ) were plated in medium alone, adherent cells counted at 24 h, and cell growth evaluated 48 h after the addition of [<sup>3</sup>H]thymidine (1  $\mu\text{Ci}$  per well) and BSA (0.1%) or bFGF (0.01, 0.1 or 1  $\text{ng ml}^{-1}$ ). Tat, bFGF, buffer or medium were added to H-UVE cells for 12–14 h and PMA (20  $\text{ng ml}^{-1}$ ) for 5–7 h. RNA was extracted and analysed by Northern blot (10  $\mu\text{g}$  for each lane)<sup>10</sup> by using a <sup>32</sup>P end-labelled oligodeoxynucleotide corresponding to the sequence +59 to +99 (5'-AACTCTTTGTCCGTTTGGG-GGCGACATCGCCGGGGAAGT-3') of collagenase IV cDNA encoding for the 72K form<sup>42</sup>, which detects a 3.4 kilobase transcript. Experiments were repeated 3–6 times.

Received 31 May; accepted 14 September 1994.

1. CDC Task Force on KS and opportunistic infections *N. Engl. J. Med.* **306**, 248–252 (1982).
2. Safai, B. et al. *Ann. Intern. Med.* **103**, 744–750 (1985).
3. Ruszczak, Z., Mayer-Da Silva, A. & Orfanos, C. E. *Am. J. Dermatol.* **9**, 388–398 (1987).
4. Regezi, J. A. et al. *Am. J. Pathol.* **143**, 240–249 (1993).
5. Ensoil, B., Barillari, G. & Gallo, R. C. *Hematol. Oncol. Clin. North Am.* **5**, 281–295 (1991).
6. Folkman, J. & Klagsbrun, M. *Science* **235**, 442–447 (1987).
7. Burgess, W. H. & Maciag, T. A. *Rev. Biochem.* **58**, 575–606 (1989).
8. Nakamura, S. et al. *Science* **242**, 426–430 (1988).
9. Saiahuddin, S. Z. et al. *Science* **242**, 430–433 (1988).
10. Ensoil, B. et al. *Science* **243**, 223–226 (1989).
11. Ensoil, B. et al. *J. clin. Invest.* (in the press).
12. Xerri, L. et al. *Am. J. Pathol.* **138**, 9–15 (1991).
13. Ensoil, B. et al. *Nature* **344**, 84–86 (1990).
14. Ensoil, B. et al. *J. Virol.* **67**, 277–287 (1993).
15. Barillari, G. et al. *J. Immun.* **149**, 3727–3734 (1992).
16. Barillari, G., Gendelman, R., Gallo, R. C. & Ensoil, B. *Proc. natn. Acad. Sci. U.S.A.* **90**, 7941–7945 (1993).
17. Albini, A., Barillari, G., Benelli, R., Gallo, R. C. & Ensoil, B. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
18. Hober, D. et al. *Clin. exp. Immun.* **78**, 329–333 (1989).
19. Fuchs, D. et al. *J. Acquir. Immune Defic. Syndr.* **5**, 424–425 (1992).
20. Fan, J., Bass, H. Z. & Fahey, J. L. *J. Immun.* **151**, 5031–5040 (1993).
21. Gay, C. G. & Winkles, J. A. *Proc. natn. Acad. Sci. U.S.A.* **88**, 296–300 (1991).
22. Okamura, K. et al. *J. Biol. Chem.* **266**, 19162–19165 (1991).
23. Cordon-Cardo, C., Vlodavsky, I., Haimovitz-Friedman, A., Hicklin, D. & Fuks, Z. *Lab. Invest.* **63**, 832–840 (1990).

24. Mahoney, S. E. et al. *J. clin. Invest.* **88**, 174–185 (1991).
25. Fawell, S. et al. *Proc. natn. Acad. Sci. U.S.A.* **31**, 664–668 (1994).
26. Hynes, R. O. *Cell* **69**, 11–24 (1992).
27. Brooks, P. C., Clark, R. A. F. & Cheresh, D. A. *Science* **264**, 569–571 (1994).
28. Buonaguro, L. et al. *J. Virol.* **66**, 7159–7167 (1992).
29. Ingber, D. J. *J. Cell Biochem.* **47**, 236–241 (1991).
30. Moscatelli, D. & Rifkin, D. B. *Biochem. biophys. Acta* **948**, 67–85 (1988).
31. Liotta, L. A., Steeg, P. S. & Stetler-Stevenson, W. G. *J. Cell Biol.* **64**, 327–336 (1991).
32. Mignatti, P., Tsuboi, R., Robbins, E. & Rifkin, D. J. *J. Cell Biol.* **108**, 671–682 (1989).
33. Albini, A. et al. *Int. J. Cancer* **83**, 775–779 (1991).
34. Seftor, R. E. B. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 1557–1561 (1992).
35. Seftor, R. E. B., Seftor, E., Stetler-Stevenson, W. G. & Hendrix, M. J. C. *Cancer Res.* **53**, 3411–3415 (1993).
36. Werb, Z., Tkemisie, I., Behrendtsen, O., Crowley, E. & Damsky, C. H. *J. Cell Biol.* **109**, 877–888 (1989).
37. Yang, J. et al. *J. Immun.* **152**, 223–226 (1994).
38. Zhang, Y.-M. et al. *Am. J. Pathol.* **144**, 51–59 (1994).
39. Popovic, M., Sargadharan, M. G., Read, E. & Gallo, R. C. *Science* **224**, 497–500 (1984).
40. Gartner, S. et al. *Science* **233**, 215–219 (1986).
41. Kubota, Y., Kleinman, H., Martin, G. R. & Lawley, T. J. *J. Cell Biol.* **107**, 1589–1598 (1988).
42. Collier, I. et al. *J. Biol. Chem.* **263**, 6579–6587 (1988).
43. Vogel, J., Hinrich, S. H., Reynolds, R. K., Luciw, P. A. & Jay, G. *Nature* **335**, 606–611 (1988).

SUPPLEMENTARY INFORMATION is available from the London Editorial Offices of Nature.

**ACKNOWLEDGEMENTS.** We thank G. Barillari (LTBC, NCI) and C. Hatch (NCI, NIH) for helpful discussion, P. Winfield (OD, NIH) for the Tat protein, P. Biberfeld (Karolinska Institute, Sweden) and P. Browning (LTBC, NCI) for some of the frozen specimens, V. Kao and A. Sethi for technical help and L. Anderson for editorial assistance.

# Site-selective oxygen isotope effect in optimally doped $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$

D. Zech\*, H. Keller\*, K. Conder†, E. Kaldis†, E. Liarokapis‡, N. Poulakis‡ & K. A. Müller\*

\* Physik-Institut der Universität Zürich, CH-8057 Zürich, Switzerland

† Laboratorium für Festkörperphysik ETH Zürich, CH-8093 Zürich, Switzerland

‡ National Technical University of Athens, Department of Physics, 157 80 Athens, Greece

IN CONVENTIONAL superconductors, the large dependence of the superconducting transition temperature ( $T_c$ ) on the isotope mass ( $T_c \propto m^{-\alpha}$ , with  $\alpha \approx 0.5$ ) gives strong evidence for electron pairing by a phonon-mediated mechanism. The copper oxide superconductors, on the other hand, exhibit a small oxygen isotope effect ( $\alpha \approx 0.02$ ; see, for example, refs 1–14), suggesting a mechanism more complex than simple phonon-mediated pairing<sup>15</sup>. To obtain further insight into this mechanism, it is important to determine the contribution from different oxygen sites in the crystal lattice to the total oxygen isotope effect. Recently, a negative isotope effect associated with oxygen atoms in the copper oxide planes of the crystal structure was reported<sup>13</sup> for highly doped  $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$ , implying that the apical/chain oxygens make a substantial contribution to the total oxygen isotope shift. Here we investigate this effect in optimally doped  $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$ , by measuring separately the contributions to the total isotope shift of the planar and apical/chain oxygen sites. The sum of the site-selective isotope shifts equals that of the total measured shift, illustrating the self-consistency of our results. In contrast to the previous work<sup>13</sup>, we find that more than 80% of the total (positive) isotope effect is associated with the copper oxide planes.

In this study the starting material was synthesized using oxides previously prepared by direct oxidation of Y, Ba and Cu metals. Sinterings at 945 °C (15 h) were repeated several times with

intermediate grindings in a glove-box. Small amounts (<2 wt%) of  $\text{BaCuO}_{2+\delta}$  as a second phase were detected by X-ray diffraction. A new two-step exchange process allowed the simultaneous preparation of samples with complete substitution at all sites ( $\langle^{16}\text{O}\rangle$  or  $\langle^{18}\text{O}\rangle$ ), and samples containing predominantly  $^{16}\text{O}$  (or  $^{18}\text{O}$ ) in the  $\text{CuO}_2$  plane sites and  $^{18}\text{O}$  (or  $^{16}\text{O}$ ) in the apical and chain sites (site-selective oxygen substitution  $\langle^{16}\text{O}_p^{18}\text{O}_{ac}\rangle$  or  $\langle^{18}\text{O}_p^{16}\text{O}_{ac}\rangle$ ). Two triplets A and B of samples, designated as  $\langle^{16}\text{O}\rangle^A$ ,  $\langle^{18}\text{O}\rangle^A$ ,  $\langle^{16}\text{O}_p^{18}\text{O}_{ac}\rangle^A$  and  $\langle^{16}\text{O}\rangle^B$ ,  $\langle^{18}\text{O}\rangle^B$ ,  $\langle^{18}\text{O}_p^{16}\text{O}_{ac}\rangle^B$ , were prepared from the same batch of the starting material. In order to exclude any influence of the thermal treatment during the isotope exchange on the oxygen isotope shift, each triplet was prepared in a simultaneously performed two-step annealing run in closed ampoules. Thus, in the first step (430 °C, 75 h) the total and at the second step (300 °C, 300 h) the site-selective oxygen exchange was performed (see, for example, refs 16, 17). For the site-selective substituted samples the gas atmosphere was changed after the temperature of the second annealing step was reached. The total oxygen content was determined using high-accuracy volumetric analysis<sup>18</sup>. The average oxygen content was  $x = 0.957(2)$  and  $0.963(3)$  for triplets A and B, respectively. The  $^{18}\text{O}$  contents in the samples were determined thermogravimetrically after  $^{18}\text{O}$  back-exchange with  $^{16}\text{O}$  (ref. 16). For the totally exchanged samples the  $^{18}\text{O}$  content was >91%. For the  $\langle^{16}\text{O}_p^{18}\text{O}_{ac}\rangle^A$  sample it could be estimated that also 15(5)% of the plane atoms were occupied by  $^{18}\text{O}$ . For the  $\langle^{18}\text{O}_p^{16}\text{O}_{ac}\rangle^B$  sample ~5% of the apex and chain positions were still occupied by  $^{18}\text{O}$ .

The site-selectivity of the oxygen exchange was then checked by Raman spectroscopy. Figure 1 shows Raman spectra of all samples with (zz) and (xx) polarizations, obtained for a 488-nm excitation. In the (zz) polarization we observe for samples  $\langle^{16}\text{O}\rangle^A$  and  $\langle^{16}\text{O}\rangle^B$  the  $A_g$  modes at 436  $\text{cm}^{-1}$  (in-phase vibration of the planar O2/O3 oxygen atoms) and at 502  $\text{cm}^{-1}$  (stretching mode of the apical oxygen O4); in the (xx) polarization the  $B_{1g}$  mode at 337  $\text{cm}^{-1}$  is seen, corresponding to the out-of-phase vibrations of the planar oxygens. In the samples  $\langle^{18}\text{O}\rangle^A$  and  $\langle^{18}\text{O}\rangle^B$ , the Raman lines are all shifted to lower frequencies in amounts consistent with the square root of the mass ratio  $^{18}\text{O}$

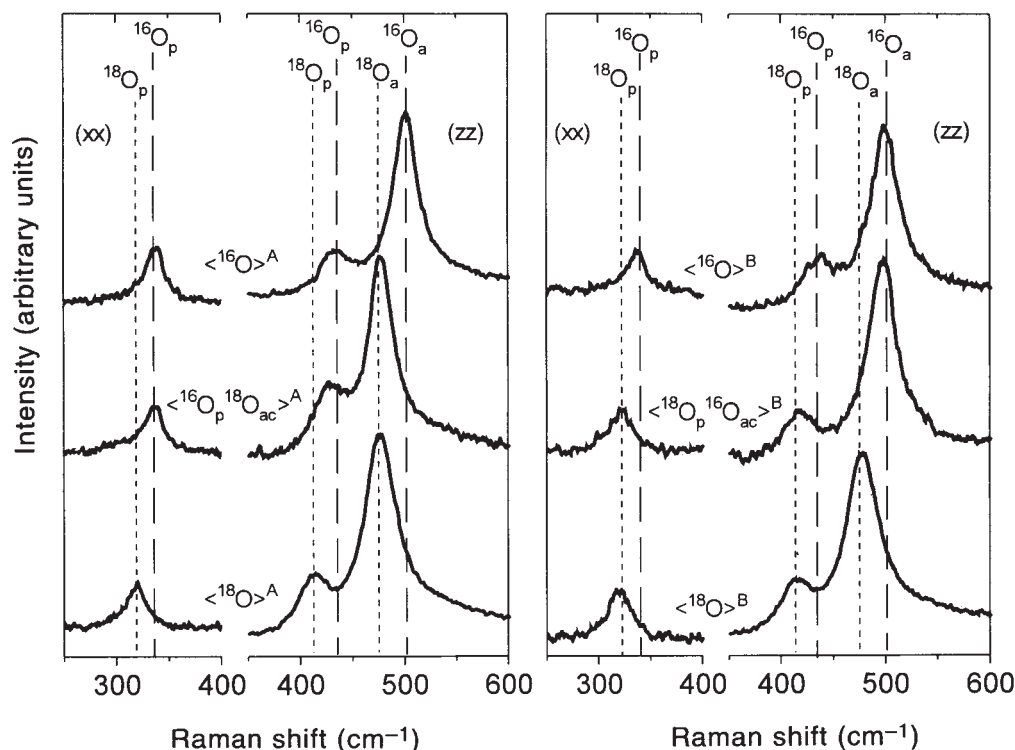


FIG. 1 Room-temperature Raman spectra of the  $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$  samples (triplet A (left) and B (right)). For (xx) polarization the line corresponds to the out-of-phase motion of the planar O2/O3 oxygens ( $^{16}\text{O}$ , 337  $\text{cm}^{-1}$ ;  $^{18}\text{O}$ , 320  $\text{cm}^{-1}$ ). For (zz) polarization the two lines correspond to the in-phase motion of the planar O2/O3 oxygens ( $^{16}\text{O}$ , 436  $\text{cm}^{-1}$ ;  $^{18}\text{O}$ , 415  $\text{cm}^{-1}$ ) and to the stretching mode of the apical oxygen O4 ( $^{16}\text{O}$ , 502  $\text{cm}^{-1}$ ;  $^{18}\text{O}$ , 478  $\text{cm}^{-1}$ ).



to  $^{16}\text{O}$ , indicating a nearly complete exchange of the  $^{16}\text{O}$  with  $^{18}\text{O}$  at the O2/O3 and O4 oxygen sites. In the site-selective sample  $\langle^{16}\text{O}_p^{18}\text{O}_{ac}\rangle^A$ , only the position of the apical oxygen O4 line is shifted to lower frequency, whereas of the two O2/O3 Raman lines only the one at  $436\text{ cm}^{-1}$  is slightly shifted to  $428\text{ cm}^{-1}$ . On the other hand, in sample  $\langle^{18}\text{O}_p^{16}\text{O}_{ac}\rangle^B$  only the two O2/O3 Raman lines are shifted, while the O4 line is not shifted apart from a small shift (498 instead of  $502\text{ cm}^{-1}$ ), probably due to the partial substitution by  $^{18}\text{O}$ . The Raman results thus confirm almost complete site-selective oxygen substitution in all samples.

To determine the oxygen isotope shift, low field (10 G, field-cooled) SQUID magnetization measurements were performed on cold-pressed pellets with identical mass and shape. The susceptibility at 10 K was found to be  $\chi \approx -0.007\text{ e.m.u. g}^{-1}$  with a variation of  $<2\%$  between all the samples. Figure 2 shows the temperature dependence of the magnetization  $M$  of all samples, normalized to  $T=10\text{ K}$ . Note that the samples of each triplet exhibit identical magnetization curves over the entire temperature range with a sharp transition and a well defined  $T_c \approx 92\text{ K}$ , indicating their excellent quality and reproducibility. In Fig. 3 the oxygen isotope shift is apparent as a small parallel shift of the magnetization curves with respect to each other. The samples  $\langle^{16}\text{O}\rangle$  and  $\langle^{18}\text{O}\rangle$  of either batch A or B show a total oxygen isotope shift of  $\Delta T_c = -0.25(5)\text{ K}$ , in agreement with previous reports (see, for example, ref. 13). Relative to the  $\langle^{16}\text{O}\rangle$  reference sample, the site-selective samples  $\langle^{16}\text{O}_p^{18}\text{O}_{ac}\rangle^A$  and  $\langle^{18}\text{O}_p^{16}\text{O}_{ac}\rangle^B$

show a partial isotope shift of  $\Delta T_c = -0.05(5)\text{ K}$  and  $-0.20(5)\text{ K}$ , respectively. Note that the two independent experiments A and B yield consistent and complementary results, clearly indicating that the main contribution ( $>80\%$ ) to the oxygen isotope shift comes from the planar oxygens. This is further supported by recent copper isotope effect studies in  $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$  by Franck *et al.*<sup>14</sup>. The isotope shift arising from the Cu sites in the  $\text{CuO}_2$  planes ( $\alpha_{\text{Cu}}$ ) is comparable to the oxygen isotope shift ( $\alpha_{\text{O}}$ ), becoming anomalously large at low doping ( $\alpha_{\text{Cu}} \approx \alpha_{\text{O}} \approx 0.9$ ). However, our results are in contradiction to the inverse site-selective oxygen isotope shift of  $\Delta T_c \approx +0.10$  to  $+0.14\text{ K}$ , reported for highly doped  $\langle^{18}\text{O}_p^{16}\text{O}_{ac}\rangle$  samples ( $x \approx 0.99$ ) by Nickel *et al.*<sup>13</sup>. Our recent magnetization studies<sup>19,20</sup> revealed that in highly doped  $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$  an intrinsic broadening and a shift of the superconducting transition occurs, strongly depending on the sample preparation parameters. As discussed in ref. 20, this makes a conclusive evaluation of the small oxygen isotope shift in highly doped  $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$  very difficult and may be a reason for the inconsistent results. Furthermore, it was found by Bornemann *et al.*<sup>12</sup> that  $\alpha_{\text{O}}$  in the related  $\text{Y}_{1-x}\text{Ca}_x\text{Ba}_2\text{Cu}_3\text{O}_8$  system remains positive independent of doping. Therefore, for this study we only used samples with excellent transition characteristics.

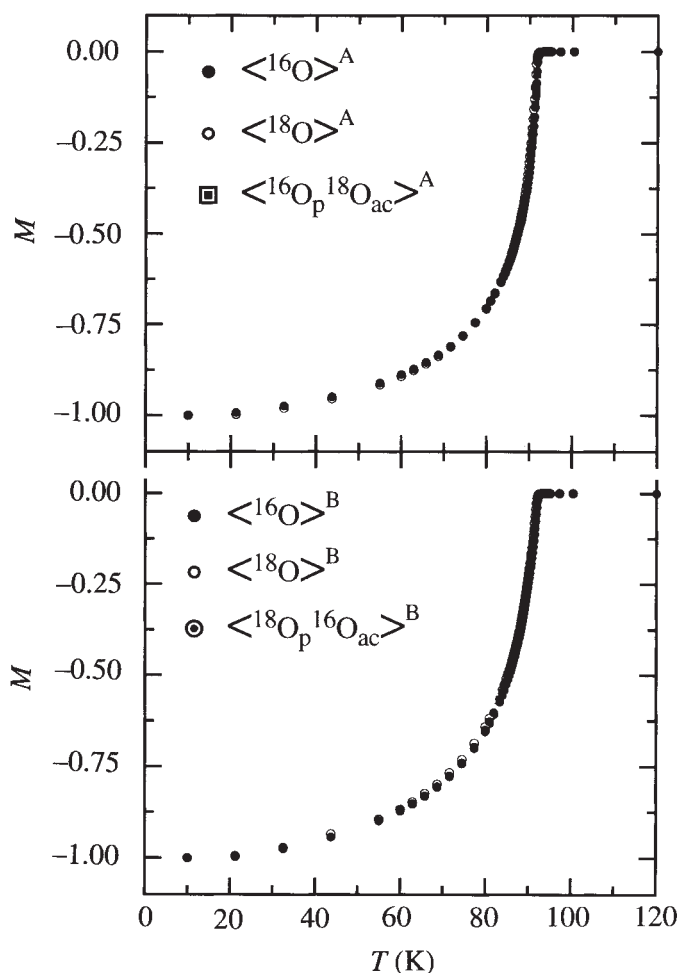


FIG. 2 Temperature dependence of the magnetization  $M$  (normalized to the value at 10 K) for the  $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$  samples (triplet A (top) and B (bottom)). Note that  $M(T)$  for the samples of each triplet overlap completely.

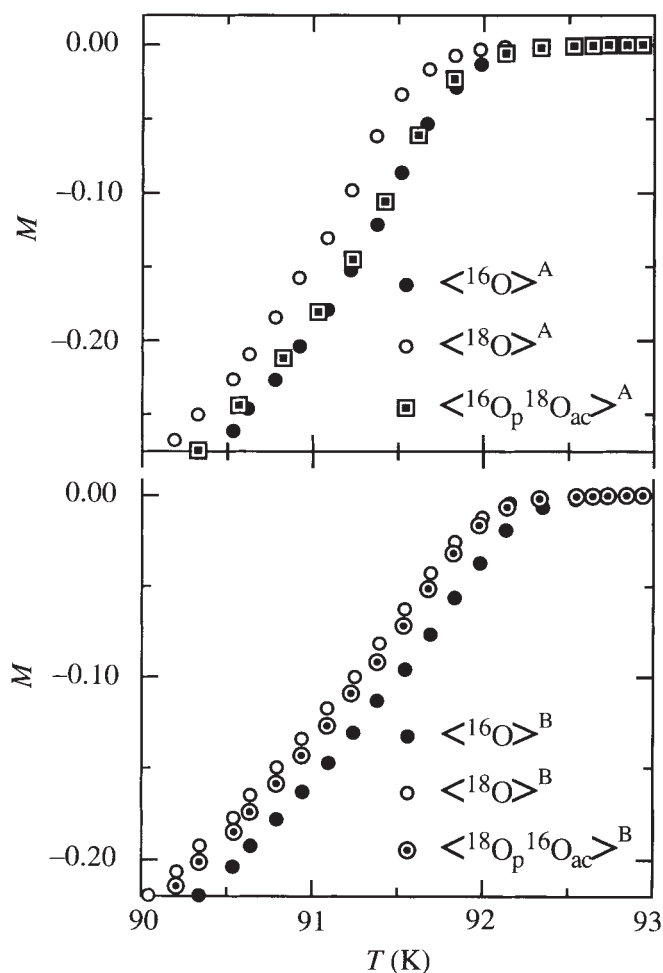


FIG. 3 Section near  $T_c$  of the normalized magnetization curves shown in Fig. 2 (triplet A, top; triplet B, bottom). Both sample pairs  $\langle^{16}\text{O}\rangle^A - \langle^{18}\text{O}\rangle^A$  and  $\langle^{16}\text{O}\rangle^B - \langle^{18}\text{O}\rangle^B$  show a total oxygen isotope shift of  $\Delta T_c = -0.25(5)\text{ K}$ . The sample  $\langle^{16}\text{O}_p^{18}\text{O}_{ac}\rangle^A$  shows almost no oxygen isotope shift  $\Delta T_c = -0.05(5)\text{ K}$  with respect to the fully  $^{16}\text{O}$  oxygenated sample, whereas the sample  $\langle^{18}\text{O}_p^{16}\text{O}_{ac}\rangle^B$  shows a partial oxygen isotope shift of about the same magnitude ( $\Delta T_c = -0.20(5)\text{ K}$ ) as the total oxygen isotope shift.

It is well established that the oxygen isotope shift coefficient  $\alpha_O$  in doped high- $T_c$  superconductors (HTSC) increases substantially with decreasing  $T_c$ , when the carrier doping is reduced from optimal levels<sup>7,12,14</sup>. In some cases  $\alpha_O$  exceeds the conventional value of 0.5 at low carrier doping, and even slightly negative values have been reported close to optimal doping<sup>9,10,12</sup>. This suggests a non-vanishing coupling between the electrons and certain high-frequency oxygen phonon modes (see, for example, refs 15, 21–23). Nevertheless, the small oxygen isotope shift, together with the high  $T_c$  for optimally doped HTSC cannot be explained using the conventional theory of Bardeen, Cooper and Schrieffer (BCS theory), and therefore a more complex phononic and/or non-phononic (polaronic/magnetic) coupling mechanism has to be considered<sup>15</sup>. As early as 1979 a group in Dubna<sup>24</sup> had already predicted that anharmonic lattice vibrations may enhance the electron–phonon coupling, causing a substantial increase of  $T_c$ . In 1987 Plakida *et al.*<sup>21</sup> proposed an anharmonic model in which the Cu–O planar tilting modes, a characteristic feature of perovskites, are considered to be important for electron–phonon pairing in HTSC. Moreover, it was suggested by one of us<sup>15</sup> that the apex oxygen motion is anharmonic and that the holes in the CuO<sub>2</sub> layers next to the apical oxygens are largely responsible for superconductivity in these materials. Indeed, several experiments (see, for example, refs 15, 21, 22), as well as computer simulations<sup>23</sup> suggest, that particular anharmonic oxygen modes are involved in the electron pairing. For instance, recent Raman investigations<sup>25</sup> on YBa<sub>2</sub>Cu<sub>3</sub>O<sub>6+x</sub> revealed that the apex oxygen mode is anharmonic. Furthermore, anharmonicity<sup>21,22</sup> may account for a large  $T_c$  and a small isotope effect coefficient  $\alpha$ , and also provides an explanation for the extremely large as well as for the small negative oxygen isotope shifts observed in doped cuprates. It is therefore not surprising that in optimally doped YBa<sub>2</sub>Cu<sub>3</sub>O<sub>6+x</sub> the contribution of the apical (chain) oxygen sites to the total oxygen isotope

shift is rather small (<20%), in view of the fact that the apex motion is anharmonic<sup>15,25</sup>.

Our results clearly demonstrate that the oxygen sites in the CuO<sub>2</sub> planes mainly contribute to the total but small oxygen isotope effect in YBa<sub>2</sub>Cu<sub>3</sub>O<sub>6+x</sub> samples close to optimal doping. This suggests that theories of electron pairing in high-temperature superconductors have to consider a phononic contribution, in which the planar tilting or buckling modes in the CuO<sub>2</sub> layers play an essential role<sup>21,22</sup>. □

Received 1 July; accepted 21 September 1994.

1. Batlogg, B. *et al.* *Phys. Rev. Lett.* **58**, 2333–2336 (1987).
2. Bourne, L. C. *et al.* *Phys. Rev. Lett.* **58**, 2337–2339 (1987).
3. Bourne, L. C., Zettl, A., Barbee, T. W. & Cohen, M. L. *Phys. Rev.* **B36**, 3990–3993 (1987).
4. Leary, K. J. *et al.* *Phys. Rev. Lett.* **59**, 1236–1239 (1987).
5. Morris, D. E., Kuroda, R. M., Markelz, A. G., Nickel, J. H. & Wei, J. Y. T. *Phys. Rev.* **B37**, 5936–5939 (1988).
6. Hoen, S. *et al.* *Phys. Rev.* **B39**, 2269–2278 (1989).
7. Crawford, M. K., Kunchur, M. N., Farneth, W. E., McCaron E. M. & Poon, S. J. *Phys. Rev.* **B41**, 282–287 (1990).
8. Franck, J. P., Jung, J., Mohamed, M. A.-K., Gyax, S. & Sproule, G. I. *Phys. Rev.* **B44**, 5318–5321 (1991).
9. Bornemann, H. J. & Morris, D. E. *Phys. Rev.* **B44**, 5322–5325 (1991).
10. Bornemann, H. J., Morris, D. E. & Liu, H. B. *Physica* **C182**, 132–136 (1991).
11. Babushkina, N. *et al.* *Physica* **C185–189**, 901–902 (1991).
12. Bornemann, H. J., Morris, D. E., Liu, H. B. & Narwankar, P. K. *Physica* **C191**, 211–218 (1992).
13. Nickel, J. H., Morris, D. E. & Ager, J. W. *Phys. Rev. Lett.* **70**, 81–84 (1993).
14. Franck, J. P., Harker, S. & Brewer, J. H. *Phys. Rev. Lett.* **71**, 283–286 (1993).
15. Müller, K. A. Z. *Phys.* **B80**, 193–201 (1990).
16. Conder, K., Kaldis, E., Maciejewski, M., Müller, K. A. & Steigmeier, E. F. *Physica* **C210**, 282–288 (1993).
17. Morris, D. E. *et al.* *Phys. Rev.* **B44**, 9556–9560 (1991).
18. Conder, K., Rusiecki, S. & Kaldis, E. *Mater. Res. Bull.* **24**, 581–587 (1989).
19. Conder, K., Krüger, Ch., Kaldis, E., Zech, D. & Keller, H. *Physica* **C225**, 13–20 (1994).
20. Zech, D., Keller, H., Müller, K. A., Conder, K. & Kaldis, E. in *Proc. Int. Conf. High-Temperature Superconductivity M<sup>2</sup>HTSC IV* Grenoble, France 1994; *Physica C* (in the press).
21. Plakida, N. A., Aksenov, V. L. & Drechsler, S. L. *Europhys. Lett.* **4**, 1309–1314 (1987).
22. Crespi, V. H. & Cohen, M. L. *Phys. Rev.* **B48**, 398–406 (1993).
23. Frick, M., Morgenstern, I. & von der Linden, W. Z. *Phys.* **B82**, 339–345 (1991).
24. Vujčić, G. M., Aksenov, V. L., Plakida, N. M. & Stamenković, S. *Phys. Lett.* **73A**, 439–441 (1979).
25. Ruani, G. *et al.* *Physica* **C226**, 101–105 (1994).

ACKNOWLEDGEMENTS. We thank I. Morgenstern for fruitful discussions. This work was supported by the Swiss National Science Foundation.

## Classification of chemical bonds based on topological analysis of electron localization functions

B. Silvi & A. Savin

Laboratoire de Dynamique des Interactions Moléculaires, UPR271, Université Pierre et Marie Curie, 4 Place Jussieu, 75252 Paris cédex, France

THE definitions currently used to classify chemical bonds (in terms of bond order, covalency versus ionicity and so forth) are derived from approximate theories<sup>1–3</sup> and are often imprecise. Here we outline a first step towards a more rigorous means of classification based on topological analysis of local quantum-mechanical functions related to the Pauli exclusion principle. The local maxima of these functions define ‘localization attractors’, of which there are only three basic types: bonding, non-bonding and core. Bonding attractors lie between the core attractors (which themselves surround the atomic nuclei) and characterize the shared-electron interactions. The number of bond attractors is related to the bond multiplicity. The spatial organization of localization attractors provides a basis for a well-defined classification of bonds, allowing an absolute characterization of covalency versus ionicity to be obtained from observable properties such as electron densities.

The valence theory of Lewis<sup>1</sup> remains the basis for most modern ideas on and classifications of the chemical bond<sup>2,3</sup>. Most such classifications rely on molecular-orbital and valence-bond theories within schemes involving the linear combination of atomic orbitals (LCAO)<sup>4,5</sup>.

The characterization of chemical bonds is a qualitative rather than a quantitative exercise, and the question is how one can distil the relevant information from experiment or from quantum-chemical calculations. The differential topology analysis of local scalar functions is a well-established mathematical approach that is well suited to handling this problem<sup>6</sup>. For a continuous, differentiable function  $g(\mathbf{r})$  defined for any point in three-dimensional ( $R^3$ ) space, the gradient  $\mathbf{X}$  defines a vector field. The theory of gradient vector fields has been successfully developed as a part of dynamical systems theory<sup>6,7</sup> (see, for instance, Abraham and Marsden<sup>8</sup> for a good introduction to the subject). Using this approach, one can identify trajectories of which the points corresponding to  $t \rightarrow -\infty$  and  $t \rightarrow \infty$  are respectively the  $\alpha$ - and  $\omega$ -limits. The set of  $\omega$ -limits is the set of the attractors of the dynamical system. The basin of an attractor is the set of points for which this attractor is the  $\omega$ -limit. This approach has been pioneered for chemical bonding by Bader<sup>9</sup> who emphasized the role of the electron density  $\rho(\mathbf{r})$ . Definition of attractors allows one to define basins which are recognized as atoms in molecules. Further analysis allows one to identify objects associated with bonds. Alternatively, the topological type of density domains bounded by isosurfaces can be considered. When the threshold defining the bounding isosurface is varied, the shape of a given density domain may or may not change. A change of topological type that occurs at a critical value of the threshold is called a bifurcation. The characterization of these



shapes and applications to chemical systems has been reviewed by Mezey<sup>10</sup>.

Electron density alone does not easily reveal the consequences of the Pauli exclusion principle on the bonding. The valence-shell electron-pair repulsion (VSEPR) theory<sup>11</sup> indicates that the Pauli principle is important for understanding chemical bonding. This theory has been reformulated in terms of maxima of  $-\nabla^2\rho(r)$  (ref. 12). The seminal works of Artmann<sup>13</sup>, Lennard-Jones<sup>14</sup>, and Bader and Stephens<sup>15</sup> have produced a series of localization functions which attempt to measure the Pauli repulsion by considering the Fermi hole (compare Luken and Culberson<sup>16</sup> and Becke and Edgecombe<sup>17</sup>). An alternative interpretation of these so-called electron localization functions,  $\eta(r)$ , can be made by considering the excess local kinetic energy due to Pauli repulsion<sup>18</sup>. The local kinetic energy  $K(r)$  is:

$$K(r) = -\frac{1}{2} \sum_{i=1}^N \int \Psi^* \nabla_i^2 \Psi \, d\tau'$$

where  $\Psi$  is an  $N$ -particle wavefunction and the prime indicates that the integration is performed over the space and spin coordinates of all particles but one.

Consider a system of fermions and a system of bosons with identical densities. The ground-state local kinetic energy of the non-interacting bosonic system is a lower bound to the local kinetic energy of the fermionic one<sup>19</sup>. The excess local kinetic energy due to the Pauli principle is just the difference between the two. Where electrons are alone or form pairs of opposite spins, the Pauli principle has little influence on their behaviour and they almost behave like bosons. In such regions the excess local kinetic energy has a low value. For practical purposes, we will require localization functions to have their maxima corresponding to this "bosonic" regime. The actual values of localization functions are calculated from approximate wavefunctions provided by quantum chemistry. However, it should be possible to derive procedures that will allow their determination from experimental densities<sup>20,21</sup> or from the measurement of the Wigner function<sup>22</sup>.

For gradient-type dynamical systems, zero-dimensional attractors are generic<sup>7</sup>. In differential topology, generic means typical; a generic property holds for most systems, but it can be violated in exceptional cases. Nevertheless, for examples that are relevant to chemistry, the system could belong to a continuous symmetry group which in turn implies that the attractor could

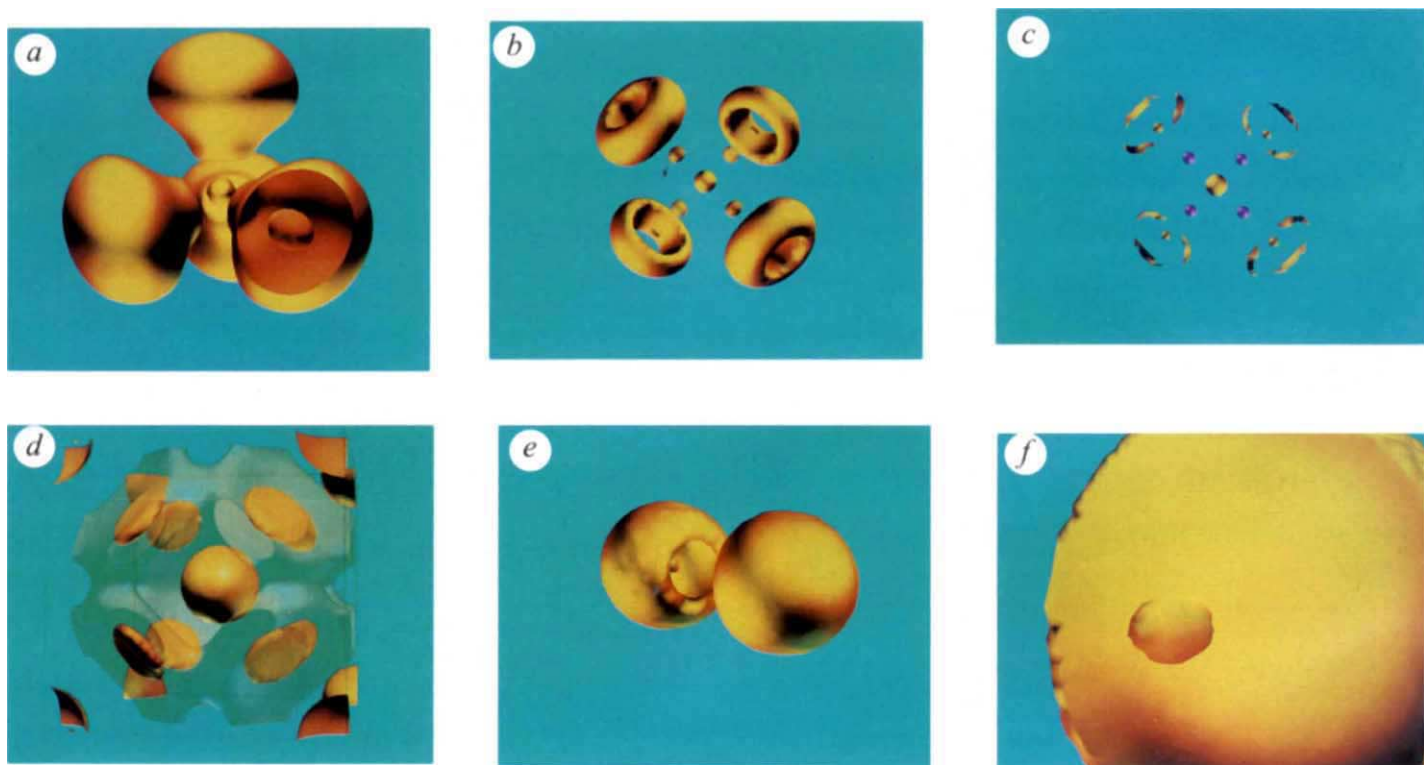


FIG. 1 Localization domains of  $\text{CF}_4$  (a–c), Li (d), LiF (e) and LiH (f). a–c, Reduction of the localization domains of  $\text{CF}_4$ . Below  $\text{ELF}=0.37$  (where ELF is the localization function; see text) there are six localization domains: five core and one valence. The bifurcation at  $\text{ELF}=0.37$  splits the common valence domain into four atomic ones. The  $\text{ELF}=0.75$  map (a) shows the carbon core surrounded by the four fluorine valence domains; the front cutting plane has been chosen so that a fluorine core domain can be seen. The bonding attractors are responsible for the bulges towards the carbon centre. A further bifurcation occurs at  $\text{ELF}=0.78$ , giving rise to bonding point attractors and non-bonding ring attractor domains as shown in b ( $\text{ELF}=0.85$ ). Each ring is itself resolved into three non-bonding point attractors for  $\text{ELF}>0.883$ . In c, the bonding attractors at which  $\text{ELF}=0.879$  are represented by purple spheres because the bounding isosurface 0.885 only encapsulates the core and non-bonding attractors. b.c.c. lithium: the core and bonding attractors are located at the  $8a$  (centre and vertices of the cubic lattice) and  $8c$

(midpoints between the centre and the vertices) positions respectively. Their domains are bounded by the  $\text{ELF}=0.625$  isosurfaces, and the  $\text{ELF}=0.575$  isosurface forms a network of channels connecting the bonding attractors. These bonding attractors are unsaturated because there are eight per cell sharing two valence electrons. For LiF e, the localization domains shown are bounded by the  $\text{ELF}=0.84$  isosurface. The fluorine valence domain (which is almost spherical at lower ELF values) shows a hole in front of the lithium core; increasing the threshold leads to a single attractor lying on the internuclear axis on the side of the fluorine core that is away from the lithium. The  $\text{ELF}=0.999$  isosurfaces of LiH (f) encapsulate, on the one hand, the lithium core and, on the other hand, a very large area which extends, in principle, to infinity. The calculation of the grid points has been limited to a box of dimensions  $7 \times 7 \times a.u.$ , therefore only one face of the largest domain can be seen. The roughness of the surface is due to interpolation limitations and a density cutoff.

be no longer zero-dimensional. We now consider some ideal localization functions in the most general case. The integral of the charge density over the basin of an attractor provides the number of electrons that belong to that basin. The Pauli principle implies that at most two electrons with opposite spins should be found in a basin in the absence of symmetry. An attractor for which  $q(A)$  is less than 2 will be called hereafter an unsaturated attractor. A "loge", in the sense of Daudel<sup>23</sup>, is the basin of a saturated attractor.

The space symmetry of atoms ( $SO(3)$  group) and of linear molecules ( $C_{\infty v}$  group), together with the Pauli principle, imply the following atomic attractors are either single point attractors located at the nucleus (K-shell attractor) or concentric spherical attractors due to the degeneracy of the off-centred attractors. The outer spherical attractor is the valence shell. In linear molecules, one can find point attractors on the internuclear axis and ring attractors centred on the axis. Spherical and ring attractors can be resolved into at least one point attractor when the symmetry is broken. Point, ring and spherical attractors are the generic topological types allowed by symmetry. They will, therefore, form the basis of our topological theory of the chemical bond.

The formation of a molecule or crystal from atoms breaks the  $SO(3)$  symmetry, and therefore leads to competition between attractors which are reorganized accordingly. To describe the chemical bond, we will first consider the  $f$ -localization domain

which we define as the set of points for which the localization function is larger or equal to  $f$ . An  $f$ -localization domain contains at least one attractor. A localization domain encapsulating more than one attractor will be called reducible, whereas an irreducible localization domain contains one (and only one) attractor. A localization domain is bounded by the isosurface  $\eta(r)=f$ . When  $f$  is varied, the localization domains undergo deformations which do not alter their topological type, if they are irreducible, whereas for reducible localization domains bifurcations occur for critical values of  $f$ . A bifurcation yields a reduction of the localization domain into several localization domains containing fewer attractors. The spatial arrangement of the irreducible localization domains is the corner-stone of our classification of chemical bonds. From a chemical point of view there are three types of attractors: core, bonding (located between the core attractors of different atoms) and non-bonding. Following Bader<sup>9</sup>, there are basically two kinds of bonding interaction: shared-electron interaction and closed-shell (or more generally, unshared-electron) interaction. Covalent, dative and metallic bonds are subclasses of the shared-electron interaction whereas ionic, hydrogen, electrostatic and van der Waals bonds belong to the other class. For shared-electron interaction there is always at least one bond attractor between the core attractors of the atoms involved in the bond. The case of hydrogen is special because it has no core attractor, and so the above classification

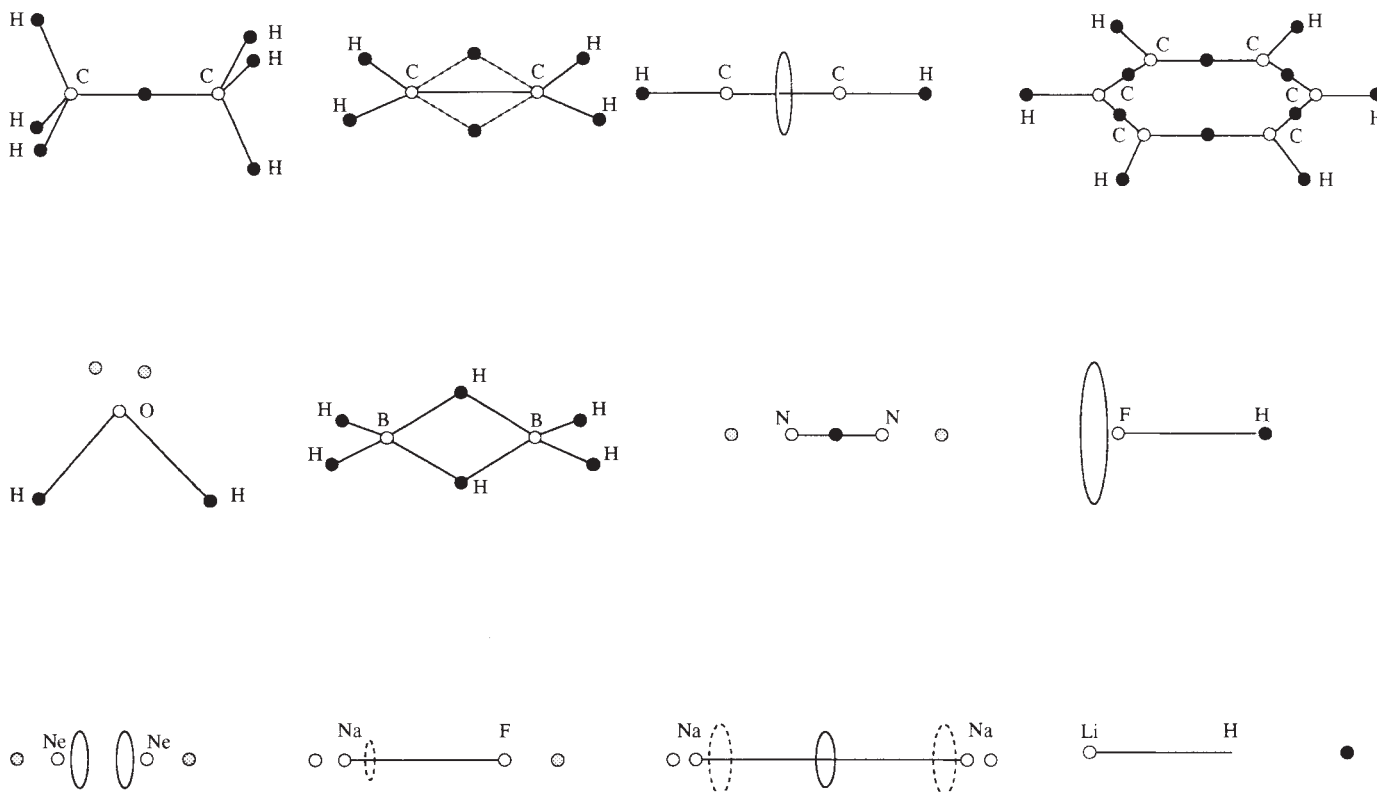


FIG. 2 Bonding diagrams of: (top row)  $C_2H_6$ ,  $C_2H_4$ ,  $C_2H_2$  and  $C_6H_6$ ; (middle row)  $H_2O$ ,  $B_2H_6$ ,  $N_2$  and  $HF$ ; (bottom row)  $Ne_2$ ,  $NaF$ ,  $Na_2$  and  $LiH$ . The core, bonding and non-bonding point attractors are presented by white, black and grey dots, respectively. Core ring attractors are drawn in dashed lines, valence attractors in solid lines. The conventional structure is given by solid lines joining the nuclei. In shared-electron interactions involving hydrogen (for example,  $C_2H_6$ ), the bond attractors are very near the protons whereas for hydrides such as  $LiH$ , the localization of the attractor is not very reliable because the localization function is very close to 1, in a domain that extends to infinity.  $C_2H_4$  and  $B_2H_6$  have very similar attractor patterns, which gives support to the

'banana' representation of the double bond in ethylene. The number of electrons related to a point attractor lying on a symmetry element may be more than two: at most, twice the order of the element. Conversely, an electron pair may give rise to a ring attractor. The triple bond attractors are degenerated onto a ring in acetylene, but correspond to a single point in  $N_2$ ; the bonding ring attractor of  $Na_2$  is related to a single pair. The diagrams for  $Na_2$  and  $NaF$  show how symmetry-breaking with respect to the free ion splits the Na L-shell sphere attractor into a point and a ring. The diagram of  $Ne_2$  corresponds to a repulsive interaction, the internuclear separation having been constrained to 1 Å.



criteria cannot be applied. However, the shape of the localization domains around the proton provides information which could be useful in this respect. When hydrogen gets a more negative character, its localization domain becomes larger (for a given value of the threshold defining the bounding isosurface).

The results used to illustrate our theory have been selected from about 100 typical systems for which the Hartree-Fock wavefunction and the ELF (ref. 17) localization function have been calculated with the *ab initio* program CRYSTAL92 (ref. 24) and visualized with the data analyser software SciAn (ref. 25). For a single determinantal wavefunction built from Hartree-Fock or Kohn Sham orbitals  $\phi_i$ ,

$$\text{ELF} = \frac{1}{1 + \left(\frac{D}{D_h}\right)^2}$$

with

$$D = \frac{1}{2} \sum_i |\nabla \phi_i|^2 - \frac{1}{8} \frac{|\nabla \rho|^2}{\rho}$$

$$D_h = \frac{3}{10} (3\pi^2)^{5/3} \rho^{5/3}$$

is calculated on a grid in the three-dimensional space. ELF has values between 0 and 1, where 1 corresponds to perfect localization. Figure 1 illustrates the idea of localization domains and shows how efficient graphic facilities can help to perform the analysis. The reduction of the localization domains of  $\text{CF}_4$  is consistent with the covalent bonding suggested by standard methods. The metallic bond is illustrated by the lithium body-centred cubic (b.c.c.) structure. Trajectories connecting the bond attractors form a three-dimensional network and pass through the saddle points. Along these paths, the value of the localization function remains nearly constant. This feature is observed for all the metallic phases we have investigated and seems to be characteristic of the metallic bond.

The attractors of 12 typical systems have been localized by a gradient search technique. The organization of these attractors is shown in Fig. 2. In both systems belonging to the shared-electron interaction set, there is always a point or ring valence attractor on the bond path whereas for the unshared-electron interaction there is nothing between the core attractors. Non-bonding lone-pair attractors are either point attractors as in  $\text{H}_2\text{O}$ ,  $\text{N}_2$  and  $\text{NaF}$ , or ring attractors in the case of  $\text{HF}$ , or both for  $\text{Ne}_2$ .

The approach presented here is an additional step towards a non-empirical basis for the classification of chemical bonds. It complements and augments the theory of Bader<sup>9</sup>. The integration of the electron density over the localization basins is in progress and should provide further information on the bond multiplicity and on the attractor saturation. □

Received 31 May; accepted 13 September 1994.

1. Lewis, G. N. *Valence and the Structure of Atoms and Molecules* (Dover, New York, 1966).
2. Kossel, A. *Ann. Physik Chemie* **49**, 229–362 (1916).
3. Langmuir, I. *J. Am. chem. Soc.* **41**, 868–934 (1919).
4. Pauling, L. *The Nature of the Chemical Bond* (Cornell Univ. Press, 1960).
5. Coulson, C. A. *Valence* (Clarendon Press, Oxford, 1952).
6. Thom, R. *Stabilité Structurale et Morphogénèse* (Interéditions, Paris, 1972).
7. Palais, J. & Smale, S. *Proc. 14th Symp. Pure Mathematics Global Analysis* (eds Chern, S. S. & Smale, S.) 223–231 (American Mathematical Soc., Providence, 1970).
8. Abraham, R. & Marsden, J. E. *Foundation of Mechanics* 507–571 (Addison-Wesley, Redwood, 1987).
9. Bader, R. F. W. *Atoms in Molecules: A Quantum Theory* (Oxford Univ. Press, 1990).
10. Mezey, P. in *Reviews in Computational Chemistry* (eds Lipkowitz, K. B. & Boyd, D. B.) Vol. 1 (VCH, New York, 1990).
11. Gillespie, R. J. *Molecular Geometry* (van Nostrand, London, 1972).
12. Bader, R. F. W., Gillespie, R. J. & McDougall, P. J. *J. Am. chem. Soc.* **110**, 7329–7336 (1988).
13. Artmann, K. Z. *Naturf.* **1**, 426–432 (1946).
14. Lennard-Jones, J. *Proc. R. Soc. A* **198**, 1–13, 14–26 (1949).
15. Bader, R. F. W. & Stephens, M. E. *J. Am. chem. Soc.* **97**, 7391–7399 (1975).
16. Luken, W. L. & Culbertson, J. C. *Int. J. quant. Chem.* **16**, 265–276 (1982).
17. Becke, A. D. & Edgecombe, K. E. *J. chem. Phys.* **92**, 5397–5403 (1990).

18. Savin, A., Jepsen, J., Andersen, O. K., Preuss, H. & von Schnering, H. G. *Angew. Chem.* **31**, 187–188 (1992).
19. Tal, Y. & Bader, R. F. W. *Int. J. quant. Chem.* **S12**, 153–168 (1978).
20. Levy, M. *Proc. natn. Acad. Sci. U.S.A.* **76**, 6062–6065 (1979).
21. Massa, L., Goldberg, M., Frishberg, C., Boehme, R. F. & La Placa, S. J. *Phys. Rev. Lett.* **55**, 622–625 (1985).
22. Royer, A. *Phys. Rev. Lett.* **55**, 2745–2748 (1985).
23. Daudel, R. *Quantum Theory of the Chemical Bond* (Reidel, Dordrecht, 1974).
24. Dovesi, R., Saunders, V. R. & Roetti, C. CRYSTAL 92 (Theoretical Chemistry Group, Univ. Turin & SERC Daresbury Laboratory, 1992).
25. Pepke, E., Murray, J., Lyons, J. & Hwu, T.-Z. *SciAn* (Supercomputer Computations Res. Inst., Florida State Univ., Tallahassee, Florida, 1993).

ACKNOWLEDGEMENTS. This work is dedicated to the memory of Linus Pauling. We thank H. G. von Schnering for his constant interest, M. Kohout for providing examples of the applicability of the ELF function, M. L. Klein for reading the manuscript, J. P. Françoise for discussions on the mathematical aspects of this work and L.-H. Jolly for computational assistance.

## Response of benthic oxygen demand to particulate organic carbon supply in the deep sea near Bermuda

F. L. Sayles, W. R. Martin & W. G. Deuser

Department of Marine Chemistry and Geochemistry,  
Woods Hole Oceanographic Institution, Woods Hole,  
Massachusetts 02543, USA

OVER the past decade an increasing body of evidence has accumulated indicating that much, perhaps most, of the deep sea floor is an environment of substantial temporal variability<sup>1–4</sup>. This variability is driven largely by seasonal changes of processes occurring in the surface waters<sup>2,3,5</sup>. The coupling of the deep sea floor environment to the surface waters is the result of rapid vertical transport of particulate matter through the water column<sup>6–8</sup>, affording only limited time for degradation before arrival at the sea floor. Studies in the Pacific Ocean have indicated that temporal variations in particulate organic carbon fluxes to the sea floor are accompanied by temporal variability in sediment oxygen demand by as much as a factor of four<sup>1,2</sup>. We report here time-series studies of oxygen fluxes into the sediments of the oligotrophic Atlantic near Bermuda which contrast sharply with these previous reports. At the Bermuda site, despite large seasonal variations in particulate organic carbon fluxes, *in situ* measured sediment oxygen consumption does not vary significantly. These results imply that large areas of the sea floor may be characterized by seasonally invariant sediment oxygen demand.

Temporal variations in the fluxes of, and response to, particulate matter reaching the sea floor have been observed in a variety of studies. Annual variations in fluxes of up to a factor of ten have been documented in continental margin<sup>2,6,9,10</sup> and open ocean environments<sup>1,3,7,11</sup>. In virtually all cases, particle flux variability has been linked to annual cycles of productivity in the overlying surface waters. The physical and chemical characteristics of material accumulating on the sea floor also exhibit short-term temporal variability. The sudden appearance of aggregates of phytodetritus on the surface of deep-sea sediment has been documented photographically at sites in both the northeast Atlantic<sup>8,12</sup> and eastern Pacific<sup>13</sup>. These events occur on an annual cycle and have been attributed to the rapid delivery of sinking particulate matter from surface blooms<sup>8,12</sup>. The presence of pigments and other rapidly degraded organic compounds in aggregates indicates that the material has undergone relatively little degradation before reaching the sea floor<sup>4,14</sup>. Studies of the arrival and disappearance of material from a particulate pulse on the Vøring Plateau (Norwegian Sea) suggest that the response of the benthic community can occur within days<sup>15</sup>. There is also strong evidence of changes in benthic metabolic activity associ-

TABLE 1 Oxygen fluxes measured at the BATS site

| Deployment number | Start date   | Sequential day | Duration (d) | Mean oxygen flux ( $\mu\text{mol cm}^{-2} \text{d}^{-1}$ ) |
|-------------------|--------------|----------------|--------------|------------------------------------------------------------|
| 3                 | 21 Mar. 1989 | 4,002          | 18           | 0.030 ( $\pm 0.003$ )*                                     |
| 4                 | 25 May 1989  | 4,067          | 24           | 0.035†                                                     |
| 5                 | 28 Jul. 1989 | 4,131          | 25           | 0.031 ( $\pm 0.001$ )                                      |
| 12                | 4 Jun. 1991  | 4,807          | 23           | 0.034 ( $\pm 0.005$ )                                      |
| 14                | 24 Jul. 1991 | 4,857          | 28           | 0.032 ( $\pm 0.001$ )                                      |
| 15                | 2 Nov. 1991  | 4,958          | 26           | 0.031 ( $\pm 0.008$ )                                      |
| 16                | 31 Mar. 1992 | 5,108          | 29           | 0.023 ( $\pm 0.003$ )                                      |
| 17                | 26 May 1992  | 5,164          | 29           | 0.028 ( $\pm 0.001$ )                                      |
| 18                | 21 Oct. 1992 | 5,312          | 26           | 0.031 ( $\pm 0.003$ )                                      |

\* Errors are 1 standard deviation of the replicate chambers.

† Measurement based on a single chamber.

ated with variations in particulate matter fluxes. Studies on a transect of stations from the California margin to the central north Pacific have shown seasonal variations in the rate of consumption of oxygen by sediments that range from a factor of two near the margin to a factor of four in the central north Pacific<sup>1,2</sup>. A continuous sediment-trap record of particulate fluxes and periodic benthic oxygen-demand measurements on the California margin over a 2-year period document the coupling between particle input and benthic metabolic rates<sup>10</sup>.

The Bermuda Atlantic time series (BATS) site, southeast of Bermuda, is considered to be representative of a large portion of the oceans, that is, oligotrophic 'oceanic' conditions<sup>15,16</sup>. It is the location of extensive studies of biogeochemical processes in

the surface waters and of the longest and most complete series of particulate flux studies in the world ocean (in excess of 16 years)<sup>7</sup>. The site lies ~83 km southeast of the island of Bermuda in ~4,400 m of water. Primary production typically peaks in late winter/early spring with smaller peaks occasionally observed later in the year; variations over the annual cycles in 1989 and 1990 were a factor of about four (ref. 17). Particle fluxes measured at 3,200 m are closely linked to the annual cycle of primary productivity with transport from the surface to 3,200 m requiring ~30 days (ref. 7). Over the period 1978–84, variations in particle fluxes, measured over 2-month intervals, were as much as a factor of 7 and averaged a factor of 2.3 over the annual cycle<sup>7</sup>. Measured over 2-week intervals, the differences between annual high and low fluxes increase to a factor of five to six. The organic carbon flux varies similarly as the organic carbon fraction is constant at ~4–5% (ref. 7).

Our measurements of sediment oxygen consumption at the BATS site were made between March 1989 and October 1992. Oxygen fluxes into the sediment were determined *in situ* with a free vehicle benthic lander<sup>18</sup>. The lander has three independent, stirred chambers to provide replicate measurements of fluxes and is designed to land on the sea floor at a very low velocity ( $3 \text{ m min}^{-1}$ ) in order to avoid creating turbulence and the resuspension of low-density material, such as the aggregates noted above, that is likely to be the most reactive. Oxygen consumption was determined from a linear regression of concentration versus time, and the volume of the chamber solution (Fig. 1).

The most prominent feature of the sediment oxygen consumption rate data is their near constancy:  $0.030 (\pm 0.0035)$

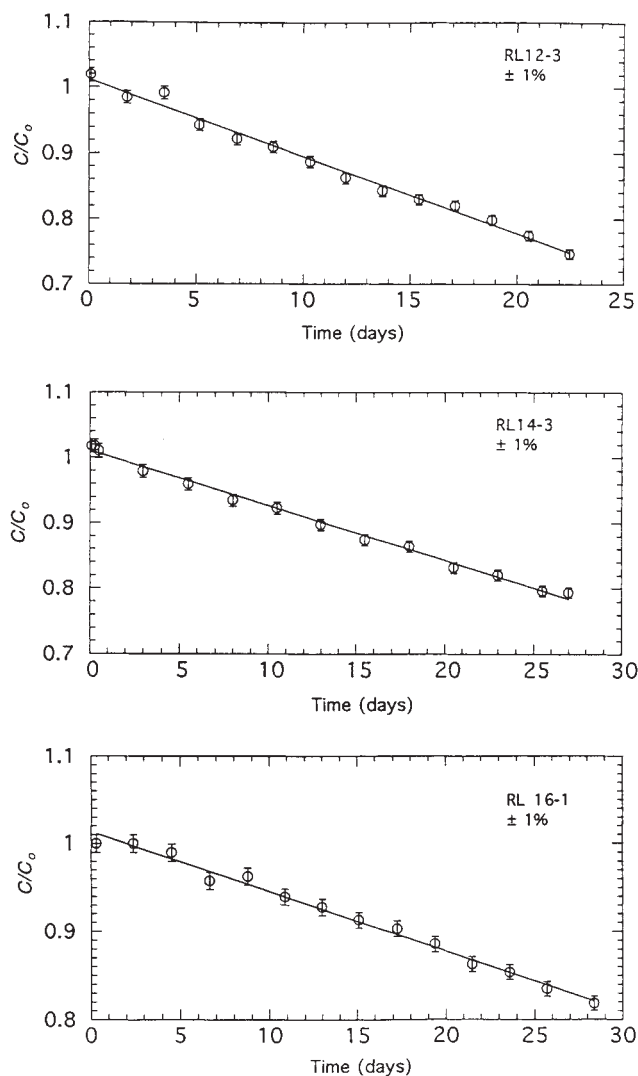


FIG. 1 Representative changes in  $\text{O}_2$  concentration with time presented (circles) as the ratio of chamber  $\text{O}_2$  ( $C$ ) to bottom water  $\text{O}_2$  ( $C_0$ ). The three panels illustrate results from three different deployments; for example, the designation RL12-3 stands for deployment 12, chamber 3. Error bars of  $\pm 1\%$  are shown. The measurements were made *in situ* throughout the duration of each experiment, 8 per deployment for the first three and 14 on each of the ensuing deployments. The measurements were made with polarographic electrodes that were standardized *in situ* relative to bottom water at the time each chamber measurement was made to compensate for electrode drift over the course of the deployments which averaged 26 days in length. Also shown (lines) is the regression of the data used, along with chamber volume, to estimate flux. The standard error in the slope is typically 4%. The uncertainty in the chamber volume was larger, being between 5 and 10%.



$\mu\text{mol cm}^{-2} \text{d}^{-1}$ ) (Table 1). Although measurements were made both during and shortly after periods of high particle flux, as well as during periods of low flux, seven of the nine  $\text{O}_2$  flux determinations fell within one standard deviation of the average value (Fig. 2a, b). During this period particle input fluctuated by a factor of about four, based on sampling every 2 weeks (Fig. 2a, b). The lack of response in oxygen consumption we have observed contrasts sharply with observations made on the continental margin of California in a similar set of experiments<sup>2</sup>. The California data document annual fluctuations in sediment oxygen consumption over a 2.5-year period (1989–91) of a factor of 1.6 and 2.6; the particulate fluxes during this period varied by factors of 3 (1989–90, monthly sampling) to  $\sim 10$  (1990–91, 10-day sampling) (Fig. 2c).

The time lag between organic carbon input and  $\text{O}_2$  flux variations depends on the rate constant for the organic carbon oxidation reaction; the amplitude of the  $\text{O}_2$  flux variation depends on both the rate constant and the amplitude of the rain rate variation<sup>19</sup>. The greater the reaction rate constant, the closer the response of oxygen consumption to rain rate variations, both in time and amplitude. We have adapted the model of ref. 19 to accept measured organic carbon rain rates and applied it to the California and Bermuda data (Fig. 2).

The analysis shows a fundamental difference between the two regions. The measured variations in the California  $\text{O}_2$  fluxes are large and strikingly similar, both in timing and magnitude, to those predicted by the model when the rate constant  $k$  for organic carbon oxidation is  $5\text{--}10 \text{ yr}^{-1}$  (Fig. 2c). The only significant departure from the model predictions is the point at about 850 days, which falls above the model curves; this departure could be explained by a rate constant that decreases as the elapsed time since an organic carbon input pulse increases. The implied rate of organic carbon oxidation is extremely rapid: for instance, the rate constant is one to two orders of magnitude greater than those inferred from high-resolution pore water profiles in the central equatorial Pacific<sup>20</sup>. In contrast, while a rate constant of  $5 \text{ yr}^{-1}$ , coupled with the organic carbon rain rate record, would require readily observable variations in  $\text{O}_2$  fluxes at the Bermuda site, measured  $\text{O}_2$  fluxes fluctuate very little (Fig. 2b). The only significant deviation from the mean flux of  $0.030 \mu\text{mol cm}^{-2} \text{d}^{-1}$  is the low value at 5,108 days, a time when the model predicts that the  $\text{O}_2$  flux should be high, not low. In order to set an upper limit on the rate constant allowed by the data, we have normalized the measured and predicted  $\text{O}_2$  fluxes to their respective average values (Fig. 3). The normalized data illustrate several points: the observed, small variations in

FIG. 2 Analysis of the coupling between the rain rate of organic carbon ( $C_{\text{org}}$ ) to the sea floor and sediment oxygen consumption at the Bermuda Atlantic time series site (a, b) and in the northeast Pacific (c, data from ref. 10). The time axis in each case is the time since the beginning of collection of the particle rain rate record. The figures show comparisons between measured  $\text{O}_2$  fluxes into the sediments (open circles) and  $\text{O}_2$  fluxes predicted by a model relating  $\text{O}_2$  consumption and organic carbon rain rate, superimposed on the record of organic carbon rain to the sea floor. The model is adapted from ref. 19. It assumes that all of the organic matter reaching the sea floor reacts at the same rate, in a reaction that is first order in the organic carbon concentration. Then, if  $I$  is the sedimentary organic carbon inventory,  $dI/dt = F(t) - kI$ , and  $I(t=0) = I_0$ , where  $F(t)$  is the rain rate of organic carbon to the sea floor (linearly interpolated at each time step from the records shown in the figure) and  $k$  is the quasi-first-order reaction rate constant.  $I_0$  was based on the average rain rate at each site; for each model run,  $I_0 = F_{\text{av}}/k$ . The effect of uncertainty in  $I_0$  is unimportant in the Bermuda data, as the particle flux record began long before the  $\text{O}_2$  flux record. In the northeast Pacific case, the effect of the uncertainty dies out rapidly because of the large  $k$  values; but the discrepancy between measured and modelled  $\text{O}_2$  flux at  $t=0$  days is due to this uncertainty. The model differential equation was solved by the fourth-order Runge–Kutta method, as implemented by ref. 23, using a time step of 1 day. In calculating the model  $\text{O}_2$  flux, it was assumed that sedimentary  $\text{O}_2$  consumption responds instantaneously to variations in organic carbon inventory<sup>19</sup>.

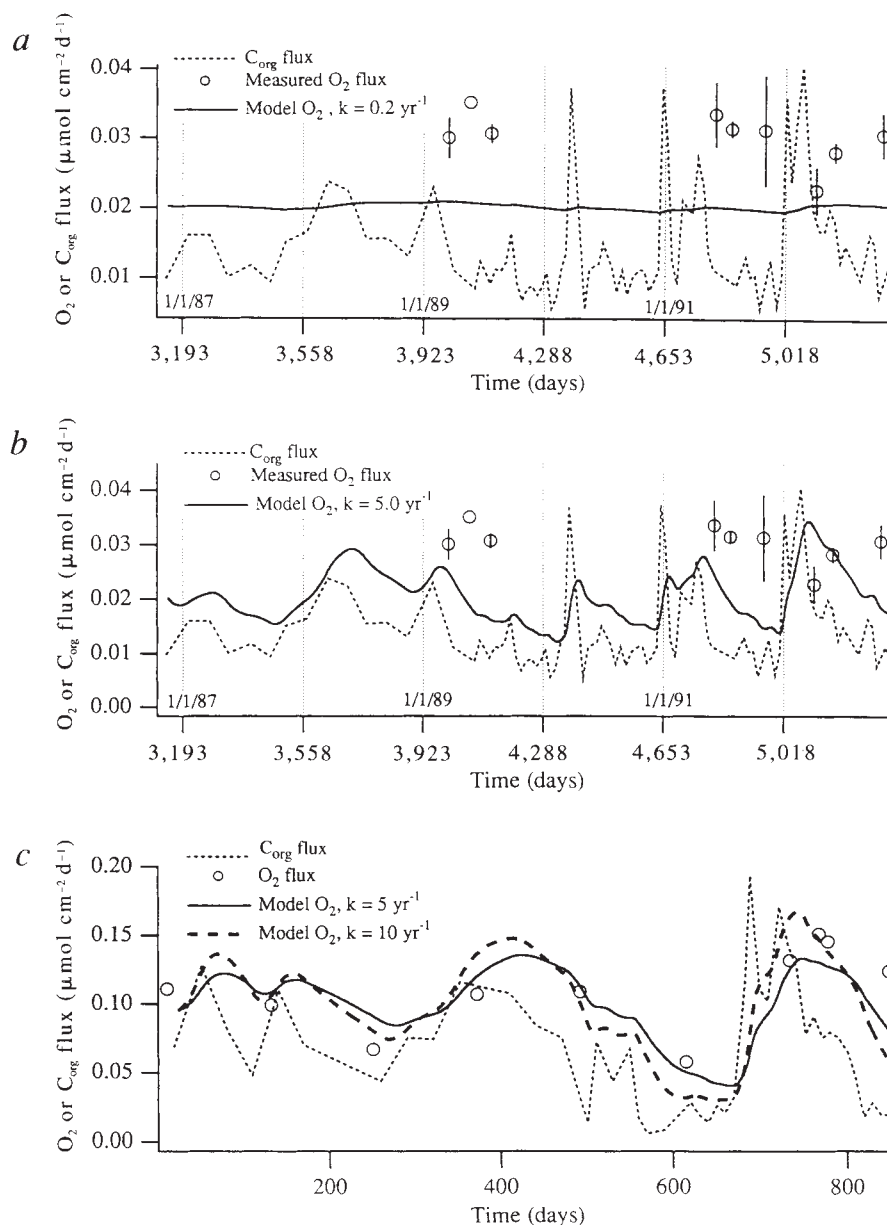
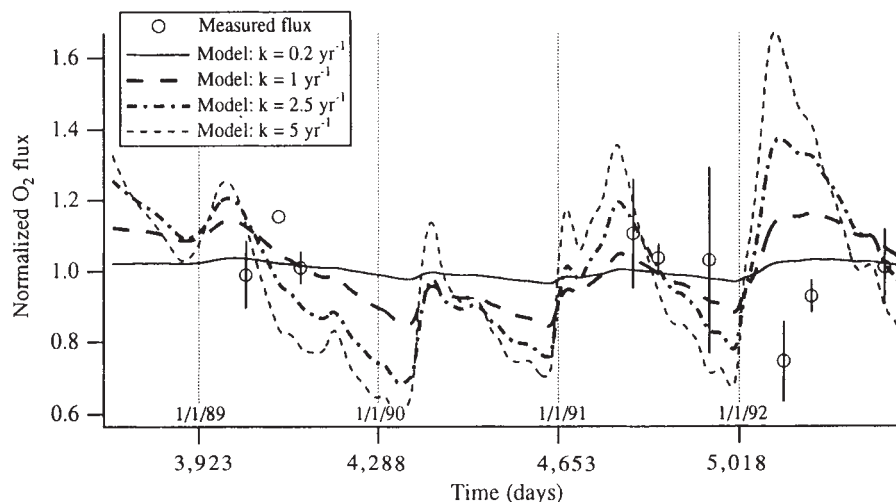


FIG. 3 The temporal variability in measured oxygen fluxes at the Bermuda site, compared to the variations predicted from the model described in Fig. 2, linking  $O_2$  fluxes to the particle rain record at that site. The measured  $O_2$  fluxes were normalized by dividing each by the average of all the measurements. Similarly, the  $O_2$  fluxes resulting from each model run were normalized by dividing by the average flux for that run.



measured  $O_2$  fluxes are out of phase with those predicted from the temporal variability in particle rain rates; with the exception of the one point mentioned above, the amplitude of the observed variations is consistent with a degradation rate constant no greater than  $1 \text{ yr}^{-1}$ ; although the variability at some time points is larger, the average standard deviation of replicate measurements ( $\sim 11\%$ ) is smaller than the predicted seasonal variability at  $k = 1 \text{ yr}^{-1}$ . No seasonal signal is apparent in the Bermuda  $O_2$  flux data; given the uncertainties in the measurements, this observation implies that the rate constant for organic carbon oxidation at this site must be smaller than  $1 \text{ yr}^{-1}$ .

The average organic carbon rain rate ( $0.0144 \mu\text{mol cm}^{-2} \text{ d}^{-1}$ ) was  $\sim 30\%$  below that required to balance the average  $O_2$  flux of  $0.030 \mu\text{mol cm}^{-2} \text{ d}^{-1}$ , based on the C:N ratio of sediment trap material (9.6) and the recently confirmed ocean-wide value for  $O_2$ :C:N stoichiometry<sup>21</sup>. It is unlikely that this discrepancy was due to artificially elevated  $O_2$  consumption within the chambers. On one deployment, one of the three chambers was deployed with a sealed bottom, so that no sediment was included; less than 2% of the  $O_2$  initially in this chamber was consumed during the 26-day experiment, compared to  $26(\pm 2)\%$  in the two chambers simultaneously deployed in the standard mode. Alternatively, measured organic carbon fluxes could be systematically low, perhaps due to consumption of organic matter within the sediment traps<sup>22</sup> (they are recovered at 2-month intervals) or to under collection of organic matter. Whatever the origin, the discrepancy does not affect our conclusions regarding the constancy of sediment oxygen fluxes at Bermuda.

The constancy of the  $O_2$  fluxes at Bermuda suggests that, if the BATS site is representative of an oligotrophic 'oceanic' regime as has been argued, the sea-floor sediments of much of the ocean undergo very limited temporal variation in metabolic rates. Our model shows that the marked contrast between Bermuda and the California margin is due primarily to the order of magnitude difference in the rate constant characterizing organic matter oxidation rather than to the substantial differences in the amplitude of temporal particle flux variations. The contrast could be due to differences in the reactivity or 'quality' of the carbon rain or to differences in biota which might govern response to a pulsed input. □

7. Deuser, W. G. *Deep-Sea Res.* **53**, 225–246 (1986).
8. Lampitt, R. S. *Deep-Sea Res.* **32**, 885–897 (1985).
9. Cole, J. J., Honjo, S. & Caraco, N. *Hydrobiology* **122**, 193–197 (1985).
10. Smith, K. L. Jr, Baldwin, R. J. & Williams, P. M. *Nature* **359**, 313–316 (1992).
11. Deuser, W. G., Muller-Karger, F. E. & Hemleben, C. *J. geophys. Res.* **93**, 6857–6862 and 6977–6978 (1988).
12. Smith, K. L. Jr, Kaufmann, R. S. & Baldwin, R. J. *Limnol. Oceanogr.* **39**(5), 1101–1118 (1994).
13. Rice, A. L. et al. *Proc. R. Soc. Edinb.* **88B**, 265–279 (1986).
14. Graf, G. *Nature* **341**, 437–439 (1989).
15. US Global Ocean Flux Study (Planning Rep. No. 5, US JGOFS Planning Office, Woods Hole Oceanographic Institution, Woods Hole, 1987).
16. US Joint Global Ocean Flux Study Long Range Plan (Planning Rep. No. 11, US JGOFS Planning Office, Woods Hole Oceanographic Institution, Woods Hole, 1990).
17. Lohrenz, S. E. et al. *Deep-Sea Res.* **39**, 1373–1391 (1992).
18. Sayles, F. L. & Dickinson, W. H. *Deep-Sea Res.* **38**, 505–529 (1991).
19. Martin, W. R. & Bender, M. L. *Am. J. Sci.* **288**, 561–574 (1988).
20. Martin, W. M., Bender, M., Leinen, M., Orcharado, J. & Isen, A. *Deep-Sea Res.* **38**, 1481–1516 (1991).
21. Anderson, L. A. & Sarmiento, J. L. *Global Biogeochem. Cycles* **8**, 65–80 (1994).
22. Gardner, W. D., Hinga, K. R. & Marra, J. J. *mar. Res.* **41**, 195–214 (1983).
23. Press, W. H., Flannery, B. P., Teukolsky, S. A. & Vetterling, W. T. *Numerical Recipes: The Art of scientific Computing* (Cambridge Univ. Press, 1986).

ACKNOWLEDGEMENTS. This study was made possible by the skill and dedication of W. Dickinson, J. Goudreau and S. Smith, as well as the Captain and crew of the RV *Weatherbird II*. This work was supported by the Chemical Oceanography Program of the US NSF.

## A line in the sea

James A. Yoder\*, Steven G. Ackleson†, Richard T. Barber‡, Pierre Flament§ & William M. Balch||

\* Graduate School of Oceanography, University of Rhode Island, South Ferry Road, Narragansett, Rhode Island 02882, USA

† Ocean Optics Program, 323 OP, Office of Naval Research, Arlington, Virginia 22217, USA

‡ Duke University Marine Laboratory, Beaufort, North Carolina 28516, USA

§ Department of Oceanography, University of Hawaii, Honolulu, Hawaii 96822, USA

|| Rosenstiel School of Marine and Atmospheric Science, 4600 Rickenbacker Causeway, Miami, Florida 33149-1098, USA

THE ocean has considerable spatial and temporal heterogeneity in biomass and productivity owing in part to the effects of ocean circulation and mixing<sup>1,2</sup>. Water mass boundaries (fronts) in coastal waters are well-known sites of enhanced biological activity<sup>3,4</sup>. Comparatively little is known of open-ocean fronts, and one of the few biological studies of an oceanic front showed phytoplankton biomass at only slightly higher densities than in surrounding waters<sup>5</sup>. Here we present photographs and measurements from satellites, aircraft, ships and the Space Shuttle Atlantis which show

Received 15 July; accepted 16 September 1994.

1. Smith, K. L. Jr & Baldwin, R. J. *Nature* **307**, 624–626 (1984).
2. Smith, K. L. Jr *Limnol. Oceanogr.* **32**, 201–220 (1987).
3. Deuser, W. G. & Ross, E. H. *Nature* **283**, 364–365 (1980).
4. Thiel, H. et al. *Biol. Oceanogr.* **6**, 203–239 (1988/1989).
5. Billett, D. S. M., Lampitt, R. S., Rice, A. L. & Mantoura, R. F. C. *Nature* **302**, 520–522 (1983).
6. Honjo, S. *Science* **218**, 883–884 (1982).



dramatic biological responses to circulation and mixing processes associated with an open-ocean front. Breaking waves (whitecaps) caused by water turbulence and mixing, and very dark green water caused by extremely high concentrations ( $>20$  mg of chlorophyll *a* per  $m^3$ ) of buoyant diatoms (*Rhizosolenia* sp.) made a distinct line in the sea visible for hundreds of kilometres. The line traced the northern edge of a westward-propagating (50 km per day) tropical instability wave (1,000-km wavelength) delineating the boundary between cold, upwelled waters and warmer waters to the north. High phytoplankton biomass and primary production associated with the extensive diatom patches may explain anecdotal observations of high animal abundance along this frontal boundary.

During the Joint Global Ocean Flux Study (JGOFS) research program in the equatorial Pacific in 1992 (refs 6, 7) a series of sea surface temperature (SST) images were obtained by the advanced very high-resolution radiometer, AVHRR, on the NOAA-11 weather satellite. These images (Fig. 1) showed westward movement ( $55\text{ km d}^{-1}$ ) of a tropical instability wave<sup>8–10</sup> near  $140^\circ\text{W}$  longitude—the site of JGOFS ship and aircraft measurements during late August and early September 1992. As illustrated in Fig. 1, tropical instability waves propagate along

the density (thermal) front delineating the boundary between cold, upwelled waters of the south equatorial current (SEC) and the warmer waters of the north equatorial countercurrent (NECC)<sup>8–10</sup>. Although the thermal front between the NEC and SECC is present year-round in the equatorial Pacific, tropical instability waves are observed in SST imagery only during the summer and autumn when the westward-flowing SEC is strongest<sup>8–10</sup>. Waves are normally observed between about  $90^\circ\text{W}$  to at least as far west as  $160^\circ\text{W}$  longitude<sup>11</sup>.

Figure 2 shows laser-induced chlorophyll fluorescence (LICF) and SST measurements acquired from a NASA P-3 research aircraft flying at 150 m altitude along  $140^\circ\text{W}$  on 21 August 1992, and LICF measurements in the same area on 25 August. LICF was obtained from the airborne oceanographic lidar (AOL) which employs a blue-green laser to excite chlorophyll *a* fluorescence at 683 nm (refs 12, 13). LICF and other radiance signals from the upper 1 m of the ocean were focused by telescope onto a photomultiplier array and after digital processing of the spectra, the strength of the LICF signal was determined at 100-m spatial resolution along the flight path of the plane. This signal is an approximate measure of chlorophyll *a* concentration and thus of phytoplankton biomass<sup>12,13</sup>. Results from 21 August show a narrow ( $<20$  km wide) spike in LICF centred near  $1.8^\circ\text{N}$  latitude coincident with an SST front separating the warm waters of the NECC from the colder, upwelled waters of the SEC (Fig. 2). Four days later on 25 August, the spike was observed near  $2.4^\circ\text{N}$  latitude, and peak LICF was now as much as 10 to 50 times higher than nearby waters (Fig. 2).

Acoustic Doppler current profiler (ADCP) measurements from the RV *Thompson* on 24–26 August showed the east–west current component on both sides of the front was to the west at  $80\text{--}90\text{ cm s}^{-1}$ ; the north–south component was strongly convergent at up to  $40\text{ cm s}^{-1}$  in the upper 100 m of the water

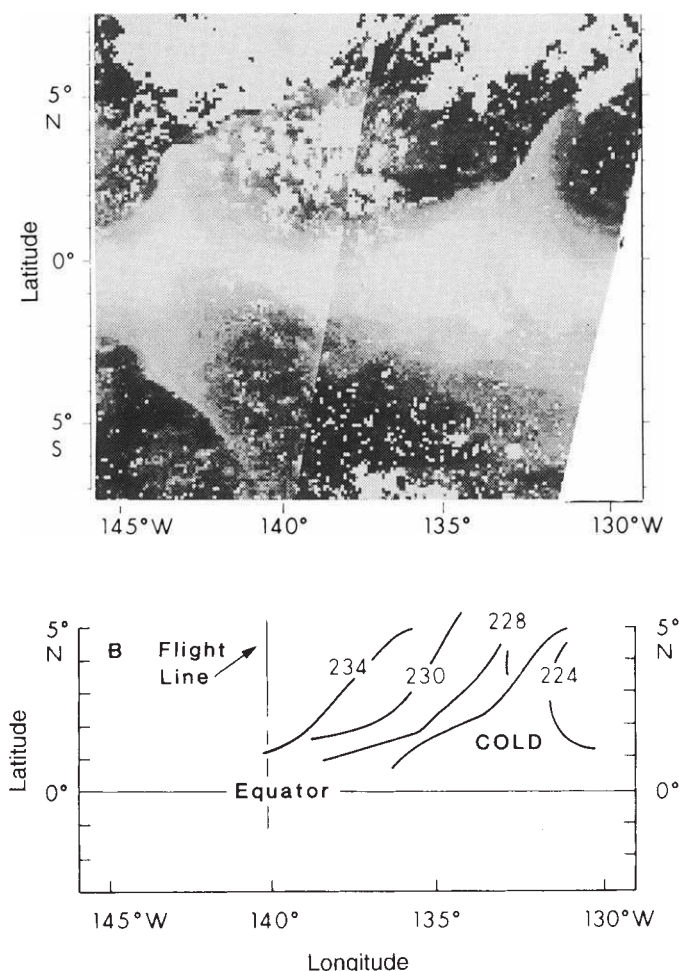


FIG. 1 Top, AVHRR-SST composite image from 11 August 1992 (day 224 of the year), showing surface temperature in the equatorial Pacific. Lightest shades of grey are cold ( $\sim 24^\circ\text{C}$ ) upwelled waters (centred around the Equator) of the south equatorial current (SEC), whereas darkest shades are warmer waters ( $\sim 28^\circ\text{C}$ ) of the north equatorial countercurrent (NECC). Clouds are white. Note the tropical instability wave (upper right) along the boundary between the SEC and the NECC. Bottom, Drawing based on a 10-d time series (224–234) of AVHRR-SST images showing the westward progression of the wave. The flight line of the NASA P-3 aircraft on 21 August 1992 (234) is also shown.

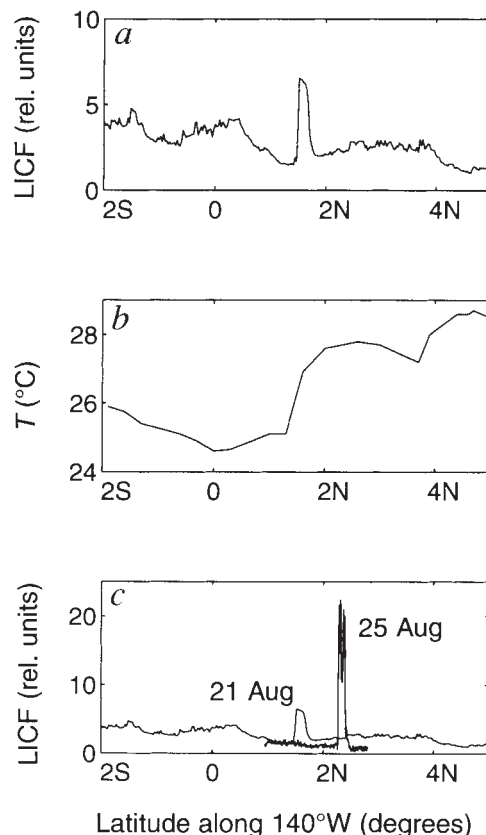


FIG. 2 a, Laser-induced chlorophyll fluorescence (LICF), and b, surface temperature ( $T$ ) along  $140^\circ\text{W}$  on 21 August 1992. c, LICF on 21 and 25 August 1992.

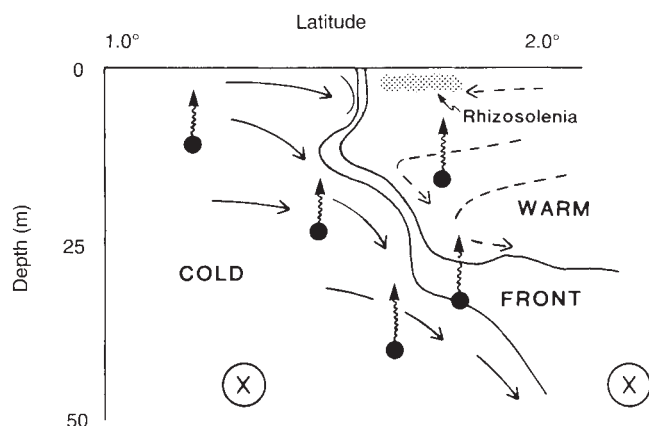


FIG. 3 Schematic drawing illustrating a possible mechanism for concentrating buoyant *Rhizosolenia* (filled circles) near the convergence. Solid and dashed arrows, northward and southward current components, respectively; solid lines, frontal boundary separating warm and cold waters; 'X' a westward component to surface flow. In this model, buoyant cells could grow on either side of the convergent front. However, cells rise to the surface and accumulate on the warm side of the front, where calculations based on ADCP measurements showed that downwelling velocities are weak compared to the strong downwelling ( $1 \text{ cm s}^{-1}$ ) on the cold side.

column at the frontal boundary<sup>14</sup>. Within the 2-km-wide frontal zone, calculations based on ADCP and hydrographic measurements indicate that a 70-m-deep slab of cold ( $24\text{--}25^\circ\text{C}$ ) water was subducting (at downwelling velocities of up to  $1 \text{ cm s}^{-1}$ ) and moving to the north beneath a 40-m-thick slab of warmer water<sup>14</sup>.

Measurements on water samples collected within the frontal region on 25 August by the RV *Thompson* showed that dense aggregations (patches) of the buoyant diatom, *Rhizosolenia* sp. (tentatively identified as *R. castracanei*, Fig. 4e) caused the high LICF. The patches were 1–2 m deep and located within the relatively warm waters in the frontal zone. Chlorophyll *a* concentrations of the patches ranged from 5 to  $29 \text{ mg m}^{-3}$ , compared to typical concentrations in nearby waters of only  $0.3 \text{ mg m}^{-3}$ . Photosynthetic index from water samples collected from the patches averaged 68 mg C per mg chlorophyll *a* per day, indicating that the diatom cells were physiologically active and growing relatively fast ( $\sim$  one cell division per day, assuming a diatom C:Chl *a* ratio of 60:1).

Figure 3 illustrates a possible mechanism for concentrating buoyant organisms at the convergence to form the observed *Rhizosolenia* sp. patches. This model is consistent with the ADCP results in that buoyant *Rhizosolenia* accumulated in the relatively warm water, where downwelling velocities were low compared to the colder water on the other side of the frontal boundary. An implication of this simple model is that *Rhizosolenia* may be taking up nutrients and growing most rapidly in the colder, upwelled waters, although it occurs at highest abundance in the warmer waters on the other side of the convergence. Previous studies show that some species of *Rhizosolenia* adjust their buoyancy and migrate vertically between surface waters and nutrient-rich waters deeper in the water column: Villareal *et al.*<sup>15</sup> report ascent rates for *Rhizosolenia* spp. as high as  $6.4 \text{ m h}^{-1}$ .

Photographs taken from viewing perspectives covering an  $\sim 10^\circ$  range in scale illustrate the spectacular effects of this feature. The photograph on the cover shows a line tracing the edge of a tropical instability wave which, given its location, size and shape, is probably the same wave observed 4 days later in SST (Fig. 1). A higher-resolution photograph (Fig. 4a) from the Space Shuttle Atlantis shows the apparent vertical relief of the feature. Photographs from the aircraft and ship (Fig. 4b–d) show lines (patches) of dark-green water adjacent to the convergence, as well as breaking waves (whitecaps) localized in a narrow zone at the convergence. Based on close examination of a video taken from the cockpit of the P-3 as it flew over the convergence and associated *Rhizosolenia* sp. patches on 25 August, the apparent vertical relief of the line is caused by the apparent increase in sea level owing to the steep, breaking waves (whitecaps) localized within a narrow zone at the convergence. Backshading provided by the heavily pigmented diatom patches

accentuates the whitecaps and contributes to the illusion of a rapid change in mean sea level.

NASA astronauts also photographed similar features in the tropical Pacific, between  $2^\circ$  and  $7^\circ \text{ N}$  latitude and  $105^\circ$  to  $170^\circ \text{ W}$  longitude, on six other occasions since 1984. All of these observations were made during the late boreal summer, autumn, or early winter (August to the following January). The association between instability waves and lines in the sea may be more than coincidental. Strong horizontal and vertical current shear and strong convergence of surface currents are associated with instability waves (refs 6, 7, 11 and P.F., R. Knox and P. Niiler, unpublished data). As illustrated in Fig. 3, strong convergence at the NECC and SECC boundary is a necessary condition for forming surface patches of the buoyant *Rhizosolenia* sp. The requirement for strong convergence may explain why visible fronts (caused by breaking waves and *Rhizosolenia* sp. patches) are observed in Space Shuttle pictures only during the instability wave season and within the latitude and longitude boundaries where the waves are common.

Extremely high concentrations and high primary productivity of *Rhizosolenia* at the NEC and SECC convergence may be an important element of a localized and productive food web, which may partially explain anecdotal observations of very high animal abundance there. Reporting from  $2.5^\circ \text{ N}$ ,  $85^\circ \text{ W}$  (which is east of where tropical instability waves are commonly observed, but where the NEC/SECC front is still well-developed<sup>8–10</sup>), while drifting along a “distinct line” of whitecaps and flotsam delineating “the meeting place of the great ocean currents”, Beebe observed:

“... here was a concentration of organisms greater than I have ever seen—the larger dotting the water and making visible its depths, the minute so abundant that in places they were of the consistency of soup.”<sup>16</sup> Another study<sup>17</sup> of a strong convergence near  $2.75^\circ \text{ N}$ ,  $121^\circ \text{ W}$ , and having similar hydrographic characteristics to the one we observed, reports “High biological activity was associated with the front; ... most of the life appeared to be concentrated in the relatively narrow frontal zone.”

Recent studies of equatorial Pacific sediment cores reveal vast deposits of laminated diatom ooze formed during the Neogene by the rapid accumulation of the mat-forming diatom, *Thalassiothrix longissima*<sup>18</sup>. Most cores showing the major intervals of laminated diatom ooze were collected between  $90^\circ$  and  $130^\circ \text{ W}$  longitude and between the Equator and  $10^\circ \text{ N}$  latitude<sup>18</sup>, that is, the region where tropical instability waves are most common in the modern equatorial Pacific Ocean. Making an analogy with the *Rhizosolenia* sp. patch observed during the 1992 JGOFS study, previous investigators proposed that the “massive flux of *T. longissima* mats may represent the fall out from major frontal systems”<sup>19</sup>. Based on our results, we speculate that tropical instability waves may also have been the mechanism for accumulating *Thalassiothrix* sp. mats at the Neogene equivalent of the



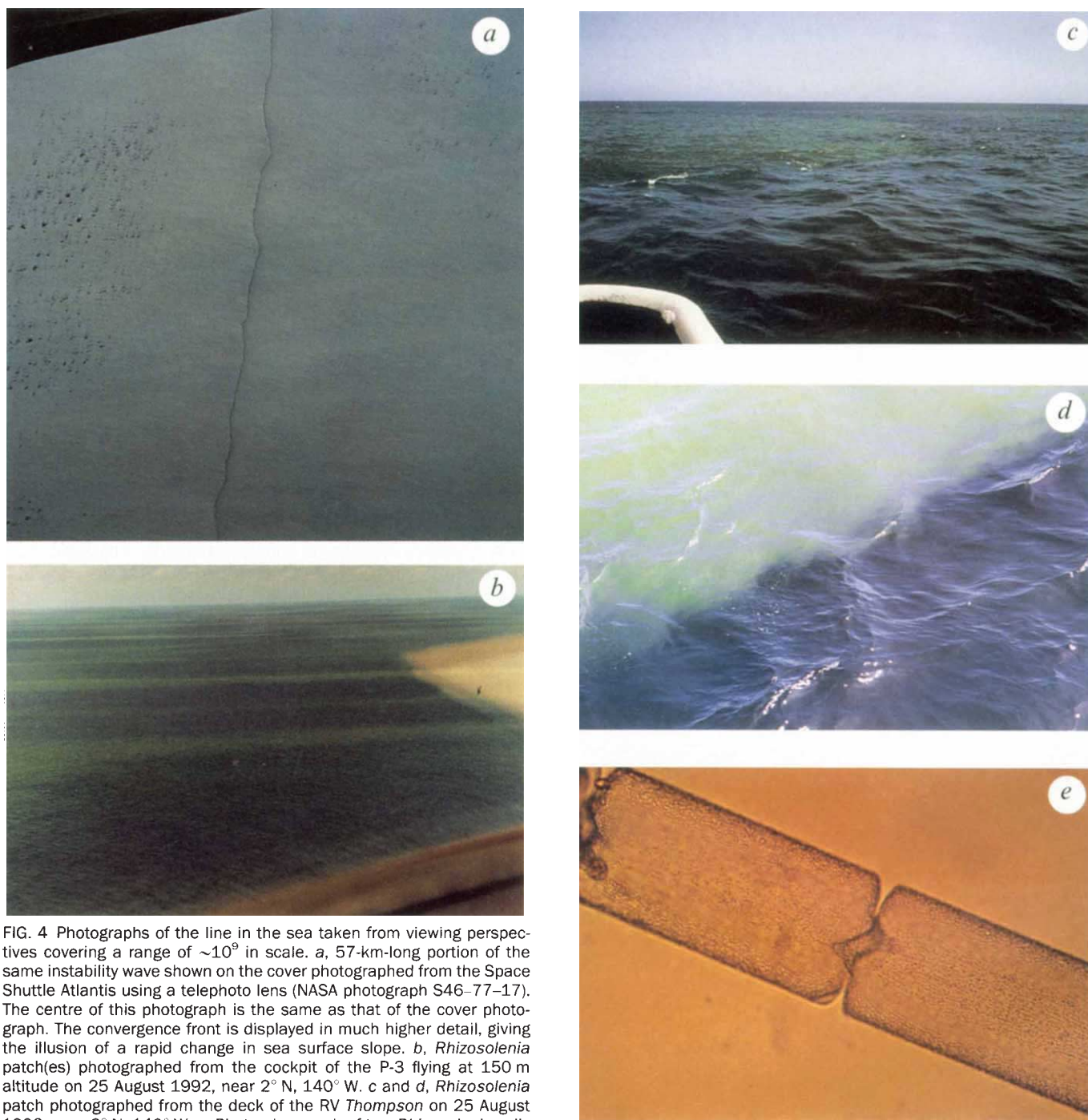


FIG. 4 Photographs of the line in the sea taken from viewing perspectives covering a range of  $\sim 10^9$  in scale. *a*, 57-km-long portion of the same instability wave shown on the cover photographed from the Space Shuttle Atlantis using a telephoto lens (NASA photograph S46-77-17). The centre of this photograph is the same as that of the cover photograph. The convergence front is displayed in much higher detail, giving the illusion of a rapid change in sea surface slope. *b*, *Rhizosolenia* patch(es) photographed from the cockpit of the P-3 flying at 150 m altitude on 25 August 1992, near  $2^\circ$  N,  $140^\circ$  W. *c* and *d*, *Rhizosolenia* patch photographed from the deck of the RV Thompson on 25 August 1992, near  $2^\circ$  N,  $140^\circ$  W. *e*, Photomicrograph of two *Rhizosolenia* cells collected from the patches illustrated above. The diameter of the cylindrically-shaped cells is  $\sim 200 \times 10^{-6}$  m.

NEC/SECC frontal boundary—a first and essential step in a sequence of physical and biological processes which eventually led to a rapid accumulation of diatom biomass in equatorial Pacific sediments<sup>18</sup>. □

Received 20 June; accepted 15 September 1994.

1. Haury, L. R., McGowan, J. A. & Wiebe, P. H. in *Spatial Pattern in Plankton Communities* (ed. Steele, J. H.) 277–328 (Plenum, New York, 1977).
2. Steele, J. H. in *The Ecology of the Seas* (eds Cushing, D. H. & Walsh, J. J.) 98–115 (Saunders, Philadelphia, 1976).
3. Yoder, J. A., Atkinson, L. P., Lee, T. N., Kim, H. H. & McClain, C. R. *Limnol. Oceanogr.* **26**, 1103–1110 (1981).
4. Marra, J., Houghton, R. W. & Garside, C. *J. mar. Res.* **48**, 851–868 (1990).
5. Fasham, M. J. R., Platt, T., Irwin, B. & Jones, K., *Prog. Oceanogr.* **14**, 129–165 (1985).
6. Murray, J. W., Barber, R. T., Bacon, M., Roman, M. R. & Feely, R. A. *Science* (in the press).
7. Barber, R. T., Murray, J. W. & McCarthy, J. J. *Ambio* **23**, 62–66 (1994).
8. Legeckis, R. *Science* **197**, 1179–1181 (1977).

9. Legeckis, R., Pichel, W. & Nesterczuk, G. *Bull. Am. met. Soc.* **64**, 133–139 (1983).
10. Miller, L., Watts, D. R. & Wimbush, M. J. *phys. Oceanogr.* **15**, 1759–1770 (1985).
11. Perigaud, C. *J. geophys. Res.* **95**, 7239–7248 (1990).
12. Hoge, F. E., Berry, R. E. & Swift, R. N. *Appl. Opt.* **25**, 39–47 (1986).
13. Smith, R. C. et al. *Appl. Opt.* **26**, 2068–2081 (1987).
14. Johnson, E. S. & Murray, J. W. *A. Meeting Oceanography Soc.* Honolulu, Hawaii, 18–22 July (1994).
15. Villareal, T., Altabet, M. & Culver-Rymsza, K. *Nature* **363**, 709–712 (1993).
16. Beebe, W. *The Arcturus Adventure* 45 (Putnam, New York, 1926).
17. Knauss, J. A. *Tellus* **9**, 234–237 (1957).
18. Kemp, A. E. S. & Baldauf, J. G. *Nature* **362**, 141–143 (1993).
19. Kemp, A. E. S., Baldauf, J. G. & Pearce, R. B. *Proc. ODP Sci. Res.* (in the press).

**ACKNOWLEDGEMENTS.** We thank the pilots, crew, and the scientific and technical personnel of the NASA-Wallops P-3 aircraft, especially P. Stegmann, R. Swift and F. Hoge. We also thank P. Stegmann, M. Lewis, S. McLean and E. Peltzer for photographs and videos, and the crew of Atlantis (STS-46) for looking for and photographing oceanographic features; E. Johnson for analyses of ADCP current measurements; C. Garside for nutrient measurements, and S. Cobb and J. Knauss for comments and suggestions. This work was supported by NASA and the US NSF.

# Diffusion in $\text{Mg}_2\text{SiO}_4$ polymorphs and chemical heterogeneity in the mantle transition zone

Daniel L. Farber<sup>\*†</sup>, Quentin Williams<sup>\*</sup>  
& Frederick J. Yerson<sup>†</sup>

<sup>\*</sup> Department of Earth Sciences and Institute of Tectonics, University of California, Santa Cruz, California 95064, USA

<sup>†</sup> Institute of Geophysics and Planetary Physics, Lawrence Livermore National Laboratory, Livermore, California 94550, USA

<sup>†</sup> Present address: Department of Earth and Space Sciences, University of California, Los Angeles, California 90024, USA

DIFFUSION in silicates plays a key role in a number of processes in the Earth's mantle, including viscous flow<sup>1–4</sup>, electrical conductance<sup>5–8</sup> and the homogenization of chemical heterogeneities. Although cation diffusion rates have been measured in olivine at high pressures<sup>9,10</sup>, no data exist on the chemical transport properties of the silicate phases thought to predominate in the transition zone of the mantle (from 400 to 700 km depth). Here we present measurements of cation diffusion in the  $\alpha$ -olivine phase and high-pressure  $\beta$ - and  $\gamma$ -spinel phases of  $\text{Mg}_2\text{SiO}_4$  at pressures up to 14 GPa. We find that diffusion rates in both high-pressure phases are about three orders of magnitude faster than that of olivine. When coupled with convective thinning, these faster diffusion rates suggest that the transition zone is more efficient at mixing chemical heterogeneities than the olivine-dominated upper mantle. Further-

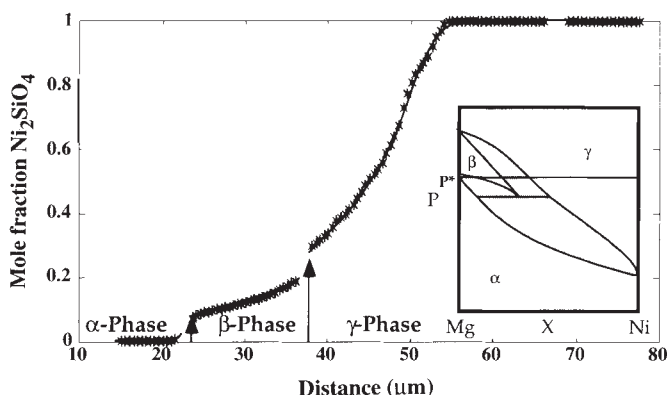


FIG. 1 Main figure, composition profile produced by Ni-Mg diffusion, as determined by electron microprobe analysis, for a sample annealed at 1,200 °C and 13.7 GPa for  $7.2 \times 10^4$  s. In all our experiments on spinel, the samples were compressed at room temperature to 5.0 GPa (above the  $\alpha$ - to  $\gamma$ - $\text{Ni}_2\text{SiO}_4$  transition), heated to 700–800 °C, and then brought to the final run pressure. The samples were then heated to 1,200 °C in 10 min. The combination of this  $P$ - $T$ - $t$  path and use of thick Ni capsules resulted in coarse ( $\sim 7$   $\mu\text{m}$  diameter) euhedral polycrystals of  $\gamma$ -phase with  $120^\circ$  grain boundaries at the start of the diffusion anneal. These grains coarsened to  $\sim 10$   $\mu\text{m}$  over the duration of the diffusion anneal. Our recovery of unfractured, unrecrystallized, millimetre-diameter single crystals of olivine with only slight evidence of undulatory extinction suggests that the diffusion anneal was carried out under a maximum differential stress of 0.1 GPa (refs 1, 28). Infrared spectroscopy demonstrated that our final olivines contained  $\sim 220$  ( $\pm 60$ ) H/( $10^6$  atoms of Si), or  $\sim 1.5$  p.p.m. by weight. We were unable to quantitatively constrain the amount of water in  $\gamma$ - $\text{Ni}_2\text{SiO}_4$ , but were able to determine that it contained less water than the co-existing olivine. Inset, schematic phase relations of the system  $\text{Ni}_2\text{SiO}_4$ - $\text{Mg}_2\text{SiO}_4$ , where  $P^*$  represents the conditions of the composition profile and  $X$  is the mole fraction of Ni.

more, we calculate that the minimum size of chemical heterogeneities in the transition zone should be of the order of metres.

We use Mg-Ni exchange to constrain divalent cation diffusion rates in olivine polymorphs for two main reasons. First, nickel has the lowest diffusion coefficient of any transition metal studied to date in olivine, and its diffusion coefficient is virtually independent of olivine composition<sup>11</sup>. Thus changes in the rate of chemical diffusion between different  $\text{Mg}_2\text{SiO}_4$  polymorphs provide a conservative estimate for the changes in the diffusion rates of iron and magnesium. Second, the behaviour of nickel is relatively insensitive to oxygen fugacity. Indeed, over the fugacity range of our experiments, nickel is expected to be entirely in the divalent oxidation state.

Single-crystal synthetic forsterite ( $\text{Mg}_2\text{SiO}_4$ ) was cut, polished and oriented, and the [001] crystallographic direction was placed perpendicular to a reservoir of powdered nickel olivine ( $\text{Ni}_2\text{SiO}_4$ ). Diffusion along this crystallographic direction is known to be about two orders of magnitude faster than in any other direction in olivine<sup>11–13</sup>, and diffusion in this direction thus controls bulk diffusion in olivine-dominated aggregates. The diffusion couples were contained in a nickel capsule with powdered NiO placed at the base of the capsule to provide both cushioning for the single crystal, and an oxygen fugacity buffer. This NiO partially reduced to metallic nickel during the experiment. The top portion of the crystal was the initial diffusion interface: in some experiments, this was dusted with platinum black to record its original position. Experiments were conducted in both a standard piston cylinder apparatus and a multi-anvil large-volume press at 1,200 ( $\pm 4$ ) °C and pressures ranging from 1 to 14 GPa. Characteristic run times were 24 hours.

At pressures below 3.3 GPa, the stable form of both  $\text{Ni}_2\text{SiO}_4$  and  $\text{Mg}_2\text{SiO}_4$  is olivine structured<sup>5,14</sup>, and the only exchange process taking place between the two phases is interdiffusion. Thus we use the equation

$$C(x, t) = 0.5(C_{+\infty} - C_{-\infty})(1 - \text{erf}[(x - x^*)/(2\sqrt{Dt})]) + C_{-\infty} \quad (1)$$

to invert for  $D$ , the diffusion coefficient, in these low-pressure, single-phase charges<sup>15</sup>. Equation (1) describes diffusion between two semi-infinite media with a concentration-independent diffusion coefficient, where  $C(x, t)$  is the concentration at position  $x$  and time  $t$ ,  $C_{+\infty}$  and  $C_{-\infty}$  are the concentrations at effectively infinite distances from the interface and  $x^*$  is the position of the interface.

At pressures above 3.3 GPa, the inversion for diffusion coefficients is complicated (and diffusion measurements in spinel facilitated) by the conversion of  $(\text{Ni}, \text{Mg})_2\text{SiO}_4$ -olivine to  $(\text{Ni}, \text{Mg})_2\text{SiO}_4$ -spinel. With increasing pressure, both the reduced solubility of nickel in olivine and enhanced solubility of magnesium in spinel produce more complex composition profiles. In diffusion profiles of this type, the composition is discontinuous between the olivine and spinel phases, with asymmetric diffusion profiles developing on either side of the boundary (Fig. 1). The lengths of the different sections of the profile reflect the value of  $D_{\text{Mg-Ni}}$  within each phase, while the magnitude of the discontinuity in composition reflects the equilibrium partitioning of nickel and magnesium between the phases, and thus represents the width of the phase loop (Fig. 1, inset).

Because of the coupled processes of diffusion and phase transformation, these types of experiments require a more sophisticated analytical treatment to invert for diffusion coefficients. The general solution for the analysis of such diffusion couples is<sup>16</sup>

$$\tilde{D}(N_2^*) = \frac{(N_2^+ - N_2^-)/V_m(N_2^*)}{2t(\delta N_2/\delta x)_{x=x^*}} \times \left[ (1 - Y^*) \int_{-\infty}^{x^*} \frac{Y}{V_m} dx + Y^* \int_{x^*}^{\infty} \frac{1 - Y}{V_m} dx \right] \quad (2)$$



where  $x^*$  is any point along the profile,  $N_2^*$  is the mole fraction of component 2 evaluated at  $x=x^*$ ,  $\bar{D}(N_2^*)$  is the binary inter-diffusion coefficient at  $x=x^*$ ,  $N_2^+(N_2^*)$  is the initial mole fraction at  $x=+\infty$  ( $x=-\infty$ ),  $V_m(N_2^*)$  is the molar volume at  $x=x^*$ ,  $t$  is the time and  $Y$  is an auxiliary variable equal to

$$\frac{N_2 - N_2^-}{N_2^+ - N_2^-}$$

As the derivation of equation (1) is general, it can be applied to diffusion in samples with more than one phase, assuming local chemical equilibrium is maintained at the phase boundaries. With this technique, the diffusion coefficient can be calculated at every point along the composition profile and there is no need to explicitly define the Boltzmann Matano interface<sup>15</sup>.

At pressures greater than ~10 GPa and less than ~14.3 GPa,  $\beta$ -(Mg, Ni)<sub>2</sub>SiO<sub>4</sub> becomes stable for compositions between  $\alpha$ -(Mg, Ni)<sub>2</sub>SiO<sub>4</sub> and  $\gamma$ -(Ni, Mg)<sub>2</sub>SiO<sub>4</sub> (Fig. 1). In these high-pressure charges, one can easily invert for  $D_{\text{Mg-Ni}}^{\beta}$  using equation (2), but because of the limited composition range of Mg-Ni solid solution in the  $\beta$ -phase structure,  $D_{\text{Mg-Ni}}^{\beta}$  is obtained from an approximation for the average diffusion coefficient in a phase with limited mutual solubility<sup>16</sup>. The growth of this  $\beta$ -(Mg, Ni)<sub>2</sub>SiO<sub>4</sub> layer is controlled by the rate of migration of the olivine- $\beta$ -spinel interface towards the Mg-rich portion of the assembly. The rate of boundary migration is proportional to the difference between  $D^{\beta\text{-phase}}$  and  $D^{\text{olivine}}$ . Thus, the rate of diffusion in the slow (olivine) phase in the diffusion couple is not the rate-controlling step in the migration of the phase boundary. Counterintuitively, the slower-diffusing phase does not dictate the rate of chemical homogenization in the charge; in fact, the slow diffusion rate in olivine increases the rate at which the phase boundary migrates.

Figure 2 shows the calculated Mg-Ni interdiffusion coefficients for the  $\alpha$ -,  $\beta$ - and  $\gamma$ -phases as a function of pressure. The most striking observation is that diffusion rates in spinel and  $\beta$ -phase are about three orders of magnitude faster than those in olivine. Indeed, although our spinel and  $\beta$ -phase results are measured within polycrystals, grain boundary effects cannot be responsible for the high diffusivity of the high-pressure phases. The grain size of both  $\beta$ -phase and  $\gamma$ -spinel in our charges was relatively coarse (~10  $\mu\text{m}$ ), and given the magnitude of their effective diffusion coefficients and our run times of the order of 10<sup>5</sup> s, we can demonstrate that the maximum overestimate of diffusion rates produced by grain boundary effects<sup>17</sup> is less than a factor of 5 (D.L.F., Q.W. and F.J.R., manuscript in preparation). Thus the observed diffusivity jumps associated with the  $\alpha$ - $\beta$  and  $\alpha$ - $\gamma$  transitions are robust, and are within

half an order of magnitude of those expected for lattice bulk diffusion alone.

We can rationalize this difference in diffusion rates on crystal-chemical grounds. Unlike olivine, the edges between divalent-cation-containing octahedra in spinel are longer than the unshared edges<sup>18</sup>. Also, the distance between neighbouring cations is shorter in spinel. Thus there is a greater likelihood of divalent cation exchange along the octahedral chains in spinels compared to those in olivine. Also, the nearly identical diffusion rates of  $\beta$ -phase and  $\gamma$ -spinel are consistent with the notable structural similarities of these phases<sup>19</sup>.

These significantly enhanced cation diffusion rates in the high-pressure phases have implications for both the rates at which subducted slabs are likely to be homogenized within the Earth's mantle, and the minimum length scale of heterogeneities in the transition zone. Using a model for mixing of chemical heterogeneities by combined diffusion and convective flow<sup>20</sup>, we calculate the rates at which divalent cations in model subducted oceanic lithosphere are homogenized (Fig. 3). This model predicts that the upper mantle should retain geochemical heterogeneities with length scales of centimetres, in accord with both previous studies and field observations<sup>20-22</sup>. But for the diffusion rates present in the transition zone, the minimum size of such heterogeneities should be of the order of metres. Figure 3 also shows the effect of convective stirring in thinning geochemical heterogeneities to length scales at which diffusive processes are able to produce homogenization.

Given this level of convective thinning, we can estimate the timescale over which Mg<sub>2</sub>SiO<sub>4</sub>-rich heterogeneities (such as the ~30-km-thick harzburgitic portion of the slab) are chemically homogenized in the transition zone. Because of the longer length scales over which cation diffusion is able to operate in the transition zone, we calculate that chemical heterogeneities should be homogenized in this region in ~75% of the time required in the upper 400 km of the Earth (Fig. 3).

If diffusional creep in the Mg<sub>2</sub>SiO<sub>4</sub> polymorphs is the dominant deformation mechanism at depth<sup>1</sup>, then the rapid diffusion rates in the  $\beta$ - and  $\gamma$ -phases suggest that the transition zone could be significantly weaker than the upper mantle. Therefore, our estimates of the difference in time for chemical equilibration between the upper mantle and transition zone could be conservative: larger strain rates could be present in this portion of the mantle. Alternatively, if the region below 400 km depth is rheologically stiff relative to the upper mantle<sup>23,24</sup>, then either a different flow mechanism or other phases (such as garnet, compare ref. 25) could be important.

Our results indicate that the transition zone is a particularly effective location for homogenization of the chemical signature

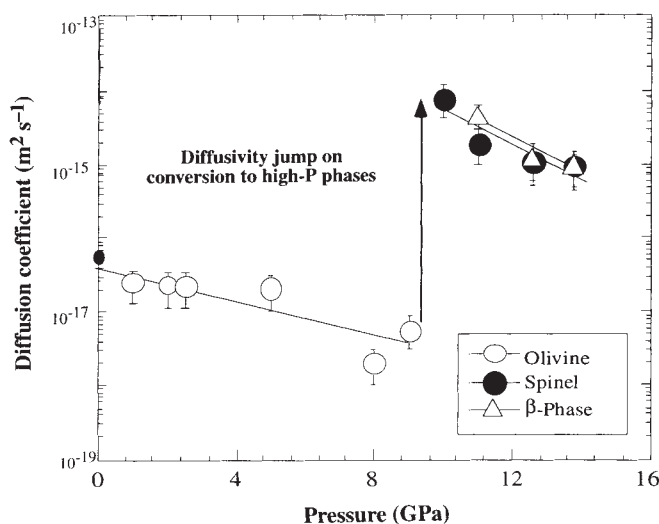


FIG. 2 Pressure dependence of Ni-Mg interdiffusion in Mg<sub>2</sub>SiO<sub>4</sub> polymorphs. The experiments below 2.5 GPa were conducted in a piston cylinder; all other data points were generated using a large-volume press. Results from the two types of apparatus near their pressures of overlap were indistinguishable. The lines are exponential fits to the data sets using a standard Arrhenius relation, which yield activation volumes for diffusion of 3.2 ( $\pm 1.2$ ), 6.6 ( $\pm 4.1$ ) and 6.7 ( $\pm 2.9$ ) cm<sup>3</sup> mol<sup>-1</sup> for olivine,  $\beta$ -phase and  $\gamma$ -spinel, respectively. The extrapolation of our olivine data to zero pressure agrees well with the ambient-pressure result of Morioka<sup>11</sup> (small filled circle).

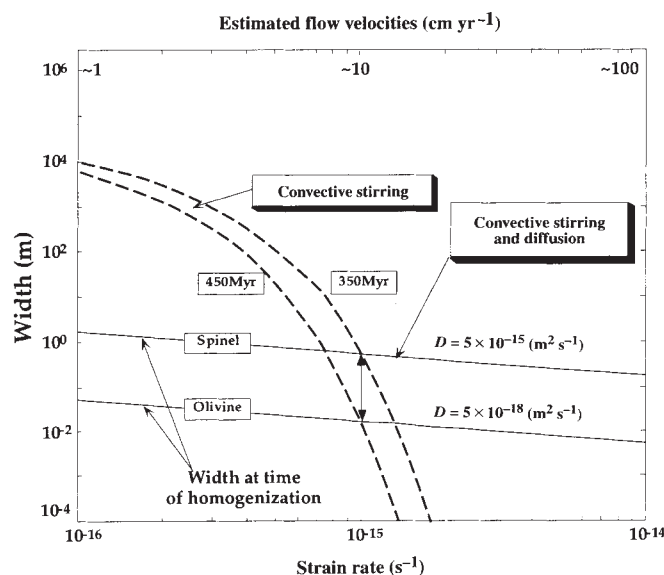


FIG. 3 Length scale of heterogeneities produced by convective stirring (dashed lines) and combined convective stirring and diffusion (solid lines). The dashed lines are isochrons showing the effect of convective stirring on thickness (as a function of strain rate) of an initially 30-km-thick sheet in chaotic flow. This sheet is designed to simulate the harzburgite layer in subducted oceanic lithosphere. Curves were calculated using an exponential thinning approximation,  $b = b_0 \exp(-\dot{\epsilon}t)$ , where  $b$  is the thickness at time  $t$ , and  $b_0$  is the initial thickness. The effective strain rates,  $\dot{\epsilon}$ , were estimated from the Kellogg and Turcotte<sup>20</sup> formulation for mixing in chaotic flow using a mean surface velocity of between  $\sim 2 \times 10^{-10}$  and  $\sim 2 \times 10^{-8} \text{ m s}^{-1}$ , and an assumed depth of convection of 700 km. The solid lines illustrate the combined effects of diffusional and convective mixing<sup>20</sup>, and represent the thickness of the sheet at the time of chemical homogenization (which is also a function of the strain rate). In this model, chemical homogenization is assumed to be achieved when the elemental concentration has fallen to 10% of its original value at the centre of the sheet. The intersection of the dashed lines and the solid lines shows the time necessary for convective stirring to reduce the thickness of the (initially) 30-km-thick sheet to the length scale at which diffusion erases the original chemical signature. We choose plate rates of  $10 \text{ cm yr}^{-1}$  to calculate the homogenization time of a 30-km-thick sheet within the upper mantle (450 Myr) and the transition zone (350 Myr). Thus for a strain rate of  $\sim 10^{-15} \text{ s}^{-1}$ , the timescales of mixing differ by a factor of 0.75, and the minimum length scales of heterogeneities by a factor of  $\sim 100$ , between the upper mantle (olivine) and the transition zone (spinel).

associated with subducted slabs. In contrast to the olivine-dominated upper mantle, the rate of migration of the boundary between the crustal portion of subducted oceanic lithosphere (garnetite) and  $\beta$ -phase or  $\gamma$ -spinel dominated assemblages within the transition zone will be largely controlled by the rapid diffusion in the  $\text{Mg}_2\text{SiO}_4$  polymorphs. Thus the transition zone is likely to be a more efficient diffusive 'slab eraser' than the upper mantle. Recent tomographic results indicate that subducted slabs may remain in the transition zone for extended periods of time<sup>26,27</sup>; therefore subduction and diffusive equilibration may, over time, enrich the transition zone with incompatible elements extracted from the shallow upper mantle. □

Received 1 March; accepted 20 September 1994.

- Karato, S. & Wu, P. *Science* **269**, 771–778 (1993).
- Karato, S., Rubie, D. & Yan, H. *J. geophys. Res.* **98**, 9761–9768 (1993).
- Jaoui, O. *J. geophys. Res.* **95**, 17631–17642 (1990).
- Karato, S. & Ogawa, M. *Phys. Earth planet. Inter.* **28**, 312–319 (1982).
- Akimoto, S. & Fujisawa, H. *J. geophys. Res.* **70**, 443–449 (1965).
- Miyamoto, M. & Takeda, H. *Nature* **303**, 602–603 (1983).
- Schock, R., Duba, A. & Shankland, T. J. *J. geophys. Res.* **94**, 5829–5839 (1989).
- Omura, K. *Phys. Earth planet. Inter.* **65**, 292–307 (1991).
- Misener, D. in *Geochemical Transport and Kinetics* (eds Hofmann, A. W., Giletti, B., Yoder, H. & Yund, R.) 117–129 (Carnegie Instn, Washington DC, 1974).

- Bertran-Alvarez, Y., Jaoui, O. & Liebermann, R. C. *Phys. Earth planet. Inter.* **70**, 102–118 (1992).
- Morioka, M. *Geochim. cosmochim. Acta* **45**, 1573–1580 (1981).
- Buening, D. & Buseck, P. J. *geophys. Res.* **78**, 6852–6862 (1973).
- Clark, A. & Long, J. in *Thomas Graham Memorial Symposium on Diffusion Processes* (eds Sherwood, J. N. et al.) 511–521 (Gordon & Breach, London, 1971).
- Katsura, T. & Ito, E. *J. geophys. Res.* **94**, 15663–15670 (1989).
- Crank, J. *The Mathematics of Diffusion* 2nd edn (Oxford Univ. Press, London, 1975).
- Wagner, C. *Acta Metall.* **17**, 99–107 (1969).
- Joesten, R. in *Diffusion, Atomic Ordering, And Mass Transport: Selected Problems in Geochemistry* (ed. Ganguly, J.) 345–395 (Springer, New York, 1991).
- Kamb, B. *Am. Miner.* **53**, 1439–1455 (1968).
- Horiuchi, H., Akaogi, M. & Sawamoto, H. in *High-Pressure Research in Geophysics* (eds Akimoto, S. & Maghnani, M. H.) 391–403 (Center for Academic Publications, Tokyo, 1982).
- Kellogg, L. & Turcotte, D. L. *J. geophys. Res.* **95**, 421–432 (1990).
- Hofmann, A. W. & Hart, S. R. *Earth planet. Sci. Lett.* **38**, 44–62 (1978).
- Hofmann, N. R. A. & McKenzie, D. *Geophys. J. R. astr. Soc.* **82**, 163–206 (1985).
- Ricard, Y., Vigny, C. & Froidevaux, C. *J. geophys. Res.* **94**, 13739–13754 (1989).
- Spada, G., Sabadini, R., Yuen, D. A. & Ricard, Y. *Geophys. J. Int.* **109**, 683–700 (1992).
- Karato, S. *Phys. Earth planet. Inter.* **55**, 234–240 (1989).
- Fukao, Y., Obayashi, M., Inoue, H. & Nenbai, M. *J. geophys. Res.* **97**, 4809–4822 (1992).
- van der Hilst, R., Engdahl, R., Spakman, W. & Nolet, G. *Nature* **353**, 37–43 (1991).
- Rubie, D., Karato, S., Yan, H. & O'Neill, H. St. *Phys. Chem. Miner.* **20**, 315–322 (1993).

ACKNOWLEDGEMENTS. We thank L. Kellogg, S. Karato, U. Christensen and J. Brenan for reviews, J. Brenan for discussions, D. Walker, M. Johnson and K. Putirka for advice on experimental techniques, C. Boro and W. Ralph for technical assistance and sample design, and J. Faust and E. Knittle for assistance with X-ray studies. We also thank the Livermore LDRD programme for the establishment of the multianvil facility. This work was supported by the IGPP, the US NSF and the Geosciences research programme of the DOE's Office of Basic Energy Science.

## High abundance of Archaea in Antarctic marine picoplankton

Edward F. DeLong, Ke Ying Wu,  
Barbara B. Prézelin & Raffael V. M. Jovine

Biology Department, Division of Ecology,  
Evolution and Marine Biology, University of California,  
Santa Barbara, Santa Barbara, California 93106, USA

ARCHAEA (archaeobacteria) constitute one of the three major evolutionary lineages of life on Earth<sup>1–3</sup>. Previously these prokaryotes were thought to predominate in only a few unusual and disparate niches, characterized by hypersaline, extremely hot, or strictly anoxic conditions<sup>4–7</sup>. Recently, novel (uncultivated) phylotypes of Archaea have been detected in coastal<sup>8</sup> and subsurface<sup>9,10</sup> marine waters, but their abundance, distribution, physiology and ecology remain largely undescribed. Here we report exceptionally high archaeal abundance in frigid marine surface waters of Antarctica. Pelagic Archaea constituted up to 34% of the prokaryotic biomass in coastal Antarctic surface waters, and they were also abundant in a variety of other cold, pelagic marine environments. Because they can make up a significant fraction of picoplankton biomass in the vast habitats encompassed by cold and deep marine waters, these pelagic Archaea represent an unexpectedly abundant component of the Earth's biota.

To investigate the distribution of pelagic marine Archaea, we concentrated picoplankton from sea water collected near Arthur Harbor, Antarctica, in late austral winter 1993. All Antarctic samples were collected at ambient temperature ( $-1.5^\circ\text{C}$ ) at the surface, or directly inside the frazil ice layer underneath the seasonal pack ice. Other marine picoplankton samples were similarly collected from a variety of sampling sites in the Northern Hemisphere (Table 1). Archaeal abundance was estimated by measuring the binding of group-specific oligonucleotide probes to ribosomal RNA extracted from the mixed picoplankton assemblage<sup>8,11–14</sup>. All Antarctic samples collected in late austral winter contained exceptionally high amounts of archaeal rRNA. About 18–30% of the total picoplankton rRNA, or 21–34% of the total prokaryotic rRNA (Table 1), was archaeal in origin. We used two probes, each of which recognizes a different region of archaeal small subunit rRNA<sup>14,15</sup> to assess archaeal abundance independently. Estimates of archaeal rRNA abundance



were essentially identical, regardless of which probe was used in quantitative experiments (Table 1). Only a small amount of archaeal rRNA (<1.0% of the total rRNA) was associated with the >1 µm cell fraction retained on prefilters, suggesting that pelagic Archaea were not endosymbionts, or associated with larger particles.

To determine the phylogenetic identity of Antarctic Archaea, we cloned and sequenced ribosomal RNA genes directly amplified from Antarctic picoplankton DNA<sup>13,16,17</sup>. For preliminary identification of the ribosomal DNA containing clones, we sequenced ~200 base pairs of the rRNA gene (positions 490–290, *Escherichia coli* numbering) in 14 randomly chosen clones. The partial rDNA sequences of all 14 clones were highly similar (80–99%) to one of two groups of pelagic marine Archaea previously detected in a variety of coastal and open ocean sea water samples<sup>8,10</sup>. Most of the clones (9 of 14) were associated with a pelagic marine archaeal group which is peripherally related to hyperthermophilic Crenarchaeota<sup>8,9</sup>. The complete sequence of the ribosomal DNA insert of a representative clone from this group (Antarctic 12) was highly similar to other pelagic marine Crenarchaeota from North American waters (97–98% unweighted sequence similarity over 908 nucleotide positions). Maximum likelihood analyses<sup>18</sup> supported a specific relationship between these Antarctic Archaea and the pelagic marine Crenarchaeota group (Fig. 1a). In addition, quantitative hybridization experiments showed that the pelagic marine Crenarchaeota rRNA accounted for 16–36% of the total extracted Antarctic archaeoplankton rRNA (Table 1).

The sequences of the remaining archaeal rDNA clones (5 of 14) were highly similar to a marine Euryarchaeota group<sup>8</sup> which appears to share a common ancestry with the aerobic moderate thermophile, *Thermoplasma acidophilum*<sup>19</sup>. The sequence of a

representative rDNA clone of this group (Antarctic 5) was highly similar to other pelagic marine Euryarchaeota (85–97% unweighted sequence similarity over 897 nucleotide positions). Maximum likelihood analyses of 792 aligned nucleotide positions of a representative clone (Antarctic 5) supported the specific relationship of these Antarctic Archaea to the planktonic marine Euryarchaeota group previously detected in coastal North American waters<sup>8</sup> (Fig. 1b).

The high archaeal rRNA abundance observed in Antarctic seas was greater than any we have yet measured in picoplankton assemblages from other marine environments (Table 1). Although low numbers of viable (but non-growing) hyperthermophilic Archaea have been found in cold surface waters<sup>20,21</sup>, the Antarctic Archaea were phylogenetically distinct from cultured hyperthermophiles, or other cultivated Archaea. Their high rRNA abundance in surface waters of Antarctica and Alaska, as well as their occurrence in cold and deep waters of temperate regions (Table 1, refs 9, 10), suggests that marine archaeoplankton are ubiquitous, active, and abundant in cold, aerobic marine waters worldwide. Because cold marine waters encompass one of the largest habitats (by volume) on Earth, these marine Archaea appear to represent a major component of the ocean's biota.

Most known Archaea thrive in habitats where, either because of the harshness of the physical environment (halophiles, hyperthermophiles) or the unusual nature of their growth substrates (methanogens), there is little (micro)biological competition. Phenotypically characterized Archaea are ecologically successful largely because of their unique physiologies and biochemistries. If this generality holds for pelagic Archaea, marine archaeoplankton may participate in unique, and considering their distribution and abundance, globally significant biogeochemical

TABLE 1 Relative binding of group-specific probes to picoplankton ribosomal RNA

| Date           | Site*  | Depth | Percentage of group-specific ribosomal RNA |             |             | Pelagic Crenarchaeota |
|----------------|--------|-------|--------------------------------------------|-------------|-------------|-----------------------|
|                |        |       | Archaea                                    | Bacteria    | Eukarya     |                       |
| Sept. 1, 1993  | Anta   | 0 m   | 26.2 (23.5)†                               | 68.6 (71.1) | 5.2 (5.4)   | 5.8 (6.0)             |
| Sept. 5, 1993  | Anta   | 0 m   | 22.7 (25.1)                                | 71.9 (69.7) | 5.3 (5.2)   | 8.2 (5.7)             |
| Sept. 9, 1993  | Antb   | 0 m   | 25.0 (30.5)                                | 63.3 (58.7) | 11.6 (10.8) | 5.3 (4.9)             |
| Sept. 13, 1993 | Antb   | 0 m   | 18.5 (24.5)                                | 68.7 (63.6) | 12.7 (11.8) | 4.3 (4.0)             |
| Jan. 18, 1994  | Alaska | 0 m   | 13.5                                       | 86.5        | NP          | NP                    |
| Aug. 29, 1992  | Pac1   | 5 m   | 0.3                                        | 71.5        | 27.7        | 0.3                   |
| Aug. 29, 1992  | Pac1   | 200 m | 4.7                                        | 72.9        | 22.3        | 2.3                   |
| Sept. 6, 1992  | Pac9   | 10 m  | ND                                         | 87.9        | 12.0        | ND                    |
| Sept. 6, 1992  | Pac9   | 110 m | ND                                         | 72.2        | 27.7        | ND                    |
| July 1, 1993   | SBB‡   | 10 m  | 2.7                                        | 97.2        | NP          | ND                    |
| July 1, 1993   | SBB    | 500 m | 18.4                                       | 81.6        | NP          | 2.0                   |
| Oct. 22, 1993  | SBBΔ21 | 15 m  | 1.2                                        | 52.6        | 46.1        | NP                    |
| Oct. 22, 1993  | SBBΔ19 | 100 m | 12.6                                       | 51.0        | 36.3        | NP                    |

Antarctic and Alaskan picoplankton were prefiltered through a 1 µm glass-fibre prefilter, and collected on 0.22 µm nylon filters. Samples SBBΔ19 and SBBΔ21 were similarly collected *in situ*, using a pump mounted on the submersible DELTA (Delta Oceanographics). Other picoplankton samples were concentrated as previously described<sup>8</sup>. Nucleic acid extraction and oligonucleotide probe hybridizations were as described<sup>8,12,13,15</sup>. Serial dilutions of nucleic acid from mixed populations or controls were applied at eight different concentrations to nylon membranes. Radioactivity was quantified by gas proportional counting and imaging (Ambis Systems, San Diego, CA). The slope of probe bound per unit rRNA was determined from 4 to 8 points in the linear portion of the binding curve. Values represent the amount of group-specific probe bound per unit rRNA, divided by the sum of archaeal + bacterial + eukaryal probe bound per unit rRNA. The percentage pelagic Crenarchaeota (see Fig. 1a)<sup>8</sup> was estimated using an oligonucleotide probe, which targets nucleotide positions 667–681 (*E. coli* numbering): CCG AGT ACC GTC TAC.

\* Sampling sites: Anta, Antarctic, 64° 46.52' S, 64° 03.27' W; Antb, Antarctic, 64° 46.77' S, 64° 04.35' W; Alaska, Resurrection Bay, Alaska; Pac1, North Pacific, 44° 02.72' N, 124° 57.30' W; Pac9, North Pacific, 31° 20.44' N, 153° 37.23' W; SBB, Santa Barbara Basin, 34° 20.44' N, 120° 04.53' W; SBBΔ19, SBBΔ21, Santa Barbara Channel, 34° 22.92' N, 120° 15.63' W.

† Numbers in parentheses are from experiments using a different archaeal-biased oligonucleotide probe, which targets nucleotide positions 363–344, *E. coli* numbering<sup>15</sup>.

‡ SBB values represent the percentage of group-specific probe bound, divided by the total prokaryotic (archaeal + bacterial) probe bound. Eukaryal rRNA abundances were not quantified in the SBB and Alaska samples.

NP, Experiment not performed; ND, not detected.

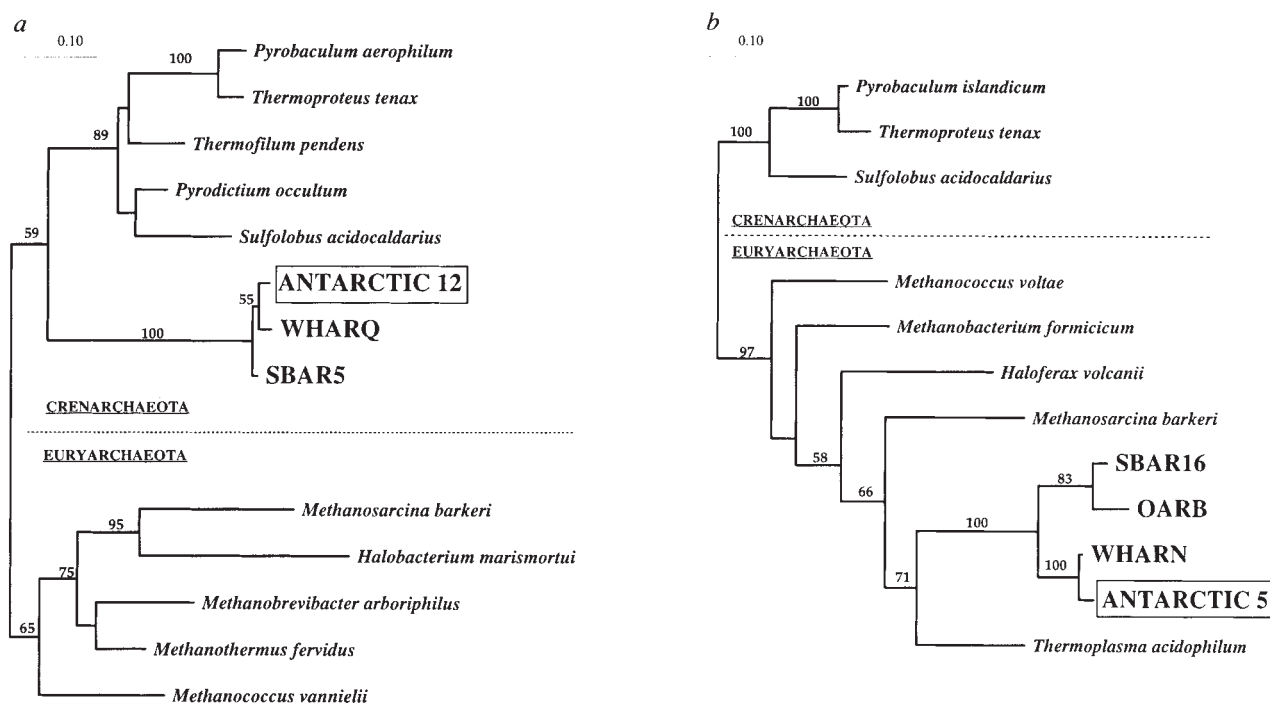


FIG. 1 Phylogenetic analysis of Antarctic Archaea. Values indicate the percentage of 500 bootstrap resamplings which support the branching pattern appearing to the right of the value. Scale bar, 10 mutations per 100 sequence positions. a, Phylogeny of a representative of the Antarctic marine Crenarchaeotal group (Antarctic 12) based on 839 nucleotide positions of rDNA. SBAR5 and WHARQ represent archaeoplankton rDNA clones, retrieved from amplified DNA of coastal North American picoplankton<sup>5</sup>. The tree was rooted using *Thermotoga maritima* as an outgroup. b, Phylogeny of a representative of the Antarctic marine Euryarchaeotal group (Antarctic 5) based on a total of 792 nucleotide positions of rDNA. SBAR16 and WHARN represent cloned rDNA genes cloned from archaeoplankton collected in coastal North America<sup>8</sup>.

processes. Because Archaea until recently were thought to be restricted to a only few disparate niches, these observations have important implications for the general ecological significance of Archaea, and their evolutionary history on Earth.

Our results, in combination with other recent reports<sup>8, 10, 22</sup> suggest that the range of phenotypes, phylotypes and environmental niches encompassed by Archaea is greater than has been previously supposed. For example, a recent study conducted in one single hydrothermal pool has greatly expanded the currently

OAR B originates from picoplankton collected in coastal surface waters off the Oregon coast. The tree was rooted using *Aquifex pyrophilus* and *Thermotoga maritima* as outgroups.

**METHODS.** Ribosomal DNA amplification, cloning and sequencing were as previously described<sup>8</sup>. Reference sequences were obtained from the Ribosomal RNA Database Project<sup>23</sup>. Phylogenetic relationships were inferred by maximum likelihood analysis<sup>17</sup> using fastDNAmI, version 1.0<sup>24</sup>. Random taxon addition and empirical base frequencies were used in the analyses. Sequences have been deposited in GenBank under the following accession numbers: OARB, U11042; ANTARCTIC 12, U11043; ANTARCTIC 5, U11044.

recognized phylogenetic diversity contained within the Crenarchaeota branch of the Archaea<sup>22</sup>. Considering that most molecular phylogenetic surveys are very preliminary and have been conducted in relatively circumscribed habitats, it seems likely that a good deal of microbial diversity remains to be discovered and described. With regard to naturally occurring prokaryotic diversity in Antarctic picoplankton populations, as well as other microbial assemblages, we may as yet be seeing only the tip of the iceberg. □

Received 28 March; accepted 22 August 1994.

- Woese, C. R. *Microbiol. Rev.* **51**, 221–271 (1987).
- Woese, C. R., Kandler, O. & Wheelis, M. L. *Proc. natn. Acad. Sci. U.S.A.* **87**, 4576–4579 (1990).
- Olsen, G. J. & Woese, C. R. *FASEB J.* **7**, 113–123 (1993).
- Whitman, W. B., Bowen, T. L. & Boone, D. R. in *The Prokaryotes* (eds Balows, A., Truper, H. G., Dworkin, M., Harder, W. & Schleifer, K. H.) 719–767 (Springer, New York, Berlin, Heidelberg, 1992).
- Jones, W. J., Nagle, D. P. & Whitman, W. B. *Microbiol. Rev.* **51**, 135–177 (1987).
- Stetter, K. O., Fiala, G., Huber, R. & Segerer, A. *FEMS Microbiol. Rev.* **75**, 117–124 (1990).
- Tindall, B. J. in *The Prokaryotes* (eds Balows, A., Truper, H. G., Dworkin, M., Harder, W. & Schleifer, K. H.) 768–808 (Springer, New York, Berlin, Heidelberg, 1992).
- DeLong, E. F. *Proc. natn. Acad. Sci. U.S.A.* **89**, 5585–5689 (1992).
- Fuhrman, J. A., McCallum, K. & Davis, A. A. *Nature* **356**, 148–149 (1992).
- Fuhrman, J. A., McCallum, K. & Davis, A. A. *Appl. env. Microbiol.* **59**, 1294–1302 (1993).
- Olsen, G. J., Lane, D. J., Giovannoni, S. J., Pace, N. R. & Stahl, D. A. *Rev. Microbiol.* **40**, 337–365 (1986).
- Stahl, D. A., Flesher, B., Mansfield, H. R. & Montgomery, L. *Appl. env. Microbiol.* **54**, 1079–1084 (1988).
- Giovannoni, S. J., Britschgi, T. B., Moyer, C. L. & Field, K. G. *Nature* **345**, 60–63 (1990).
- Stahl, D. A. & Amman, R. in *Nucleic Acid Techniques in Bacterial Systematics* (eds Stackebrandt, E. & Goodfellow, M.) 205–248 (Wiley, New York, 1991).

- Raskin, L., Stromley, J. M., Rittmann, B. E. & Stahl, D. A. *Appl. env. Microbiol.* **60**, 1232–1240 (1994).
- Sakai, R. K. et al. *Science* **230**, 487–494 (1988).
- Medlin, L., Ellwood, H. J., Stickel, S. & Sogin, M. L. *Gene* **71**, 491–499 (1988).
- Felsenstein, J. J. *molec. Evol.* **17**, 368–376 (1981).
- Woese, C. R. *Microbiol. Rev.* **58**, 1–9 (1994).
- Huber, R., Stoffers, P., Cheminee, J. L., Richnow, H. H. & Stetter, K. O. *Nature* **345**, 179–182 (1990).
- Stetter, K. O. et al. *Nature* **365**, 743–745 (1994).
- Barns, S. M., Fundyga, R. E., Jeffries, M. W. & Pace, N. R. *Proc. natn. Acad. Sci. U.S.A.* **90**, 1609–1613 (1994).
- Larsen, N. et al. *Nucleic Acids Res.* **21**, 2021–2023 (1993).
- Olsen, G. J., Matsuda, H., Hagstrom, R. & Overbeek, R. *Comp. appl. Biosci.* **10**, 41–48 (1994).

**ACKNOWLEDGEMENTS.** We thank D. Stahl for providing unpublished information on archaeal-specific oligonucleotide probes and D. Button and B. Robertson for supplying the Alaskan picoplankton sample. We thank the officers and crew of the RV *Wecoma*, the *Sproul*, and the submersible *DELTA*, and the technical support staff at Palmer Station for their cooperation and assistance. Thanks also to the RNA Database project for technical support and P. Fowler, C. Kosman, C. Preston, C. DeLaRocha, C. Gottschalk and T. J. Evens for help in the field. This work was supported by grants from the NSF, NOAA/West Coast NURC, a DuPont Young Professor grant to E.F.D., and a UCSB Faculty Career Development Award to E.F.D.



# Paternal investment inversely related to degree of extra-pair paternity in the reed bunting

Andrew Dixon, Douglas Ross, Sean L. C. O'Malley\* & Terry Burke

Department of Zoology, University of Leicester, Leicester LE1 7RH, UK

EXTRA-PAIR copulations, in which a female copulates with a male other than her mate, are known to occur in many bird species<sup>1</sup>. Here we study a wild population of reed buntings, *Emberiza schoeniclus*, using single-locus DNA fingerprinting<sup>2,3</sup> and find an exceptionally high proportion of extra-pair paternity that accounts for 55% (118/216) of young and 86% (50/58) of nests. Twelve pairs each raised two broods in a single season in which the proportion of extra-pair young differed between the two broods. A highly significant relationship between a male's parental investment and his degree of paternity was revealed by a comparison between the first and second brood of each pair, with more paternal care usually being provided at the nest that contained a lower proportion of extra-pair young. We propose that males can assess their likelihood of paternity and adjust their nestling provisioning rates accordingly.

The reed bunting is a small (about 20 g), sexually dimorphic, ground-nesting passerine, exhibiting biparental care of young. The species has an extended breeding season and pairs are capable of rearing more than one brood in a single season. Males occasionally become socially bigynous but they direct all their feeding effort to one nest, usually the primary (first hatching) nest. Nests were located and all breeding adults (except one male in 1990) and nestlings were individually marked and blood samples collected from each. DNA fingerprints (Fig. 1; Table 1) were obtained for 58 families, involving 216 chicks. Data were collected on parental provisioning behaviour using video cameras and direct observation of nests with young aged 1–8 days old (mean 9.4 h; s.d. 5.5 h observation per nest).

The fingerprinting data revealed a very high rate of extra-pair paternity (EPP): 55% (118/216) of young from 86% (50/58) of nests were not the offspring of the territorial male. Female participation in extra-pair copulations (EPCs) was virtually ubiquitous, with 97% (33/34) of females having at least one extra-pair offspring. The frequency of EPP greatly exceeded that expected from direct behavioural observation—no EPCs in the 42 observed copulation bouts. Both sexes regularly left their territory during their fertile period, so it is possible that both males and females actively seek EPCs, as has been found in other species<sup>4–6</sup>. The variance in the proportion of EPP per nest was significantly greater than expected by chance (Table 2).

The large, nonrandom variation in the extent of EPP found among reed bunting nests and the high frequency of second broods within a season provided an exceptional opportunity to investigate male parental investment in relation to actual paternity. In species exhibiting paternal care, a cuckolded male might be expected to reduce the extent of his parental care<sup>7,8</sup>. Several studies have investigated this possibility, the results of which are equivocal<sup>9–16</sup>. Unfortunately, in most of these studies paternity data are scarce. In one case in which a relationship between a male's parental effort and his degree of paternity was found it occurred in a non-monogamous system<sup>17,18</sup>, and in another the relationship was confounded by age effects<sup>15</sup>.

Among the 13 pairs that had two broods within a single breeding season, the proportion of EPP differed between the first and second broods in 12 cases. In all cases the pairs remained

TABLE 1 Minisatellite probes used in paternity analyses

| Probe    | Allele size range (kb) | N* | Heterozygosity† | Probability of genotype sharing‡ | Probability of false inclusion§ |
|----------|------------------------|----|-----------------|----------------------------------|---------------------------------|
| cEscMS1  | 4.7–9.3                | 17 | 0.94            | $7.0 \times 10^{-3}$             | 0.12                            |
| cEscMS2  | 3.1–14.0               | 26 | 0.98            | $7.9 \times 10^{-4}$             | 0.04                            |
| cEscMS7  | 1.9–6.5                | 26 | 0.95            | $4.9 \times 10^{-3}$             | 0.10                            |
| cGgaMS2  | 1.9–9.1                | 22 | 0.95            | $4.9 \times 10^{-3}$             | 0.10                            |
| Combined | —                      | —  | —               | $(1.3 \times 10^{-10})$          | $\leq 0.012$                    |

\* Number of breeding males from the study area used in this analysis.

† Heterozygosity was calculated as  $(1 - q)$ , where  $q$  is the mean allele frequency derived from the equation  $q = 1 - (1 - s)^{1/2}$ , where  $s$  is the mean band sharing frequency<sup>2</sup>.

‡ For a single locus, the probability that two unrelated individuals will share the same genotype is given by  $q^2(2 - q)$ , where  $q$  is the mean allele frequency<sup>2</sup>.

§ For a single locus, the probability of false paternal inclusion is found as  $2q - q^2$  (ref. 2). Correct maternity is assumed because conspecific nest parasitism was not observed and no maternal mismatches were detected.

|| Paternities were ascertained using at least three probes.

unchanged, nested on the same territory and had the same neighbours between each nest. Five of the 13 pairs consisted of a polygynous male and his primary female: these females retained their primary status for both their clutches and the male directed his nest-feeding effort solely to the primary female's nest in each case. The difference in the proportion of EPP between each pair's first and second nests ranged from 0 to 100%. There was no

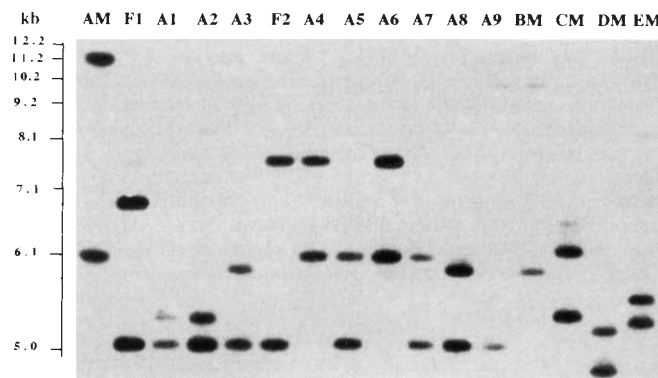


FIG. 1 Example of a single-locus minisatellite analysis of two broods of a polygynous reed bunting (AM) paired with two females (F1 and F2), using probe cEscMS1. Chicks A1 to A3 were of female F1 and were all sired by extra-pair males. Chicks A4 to A9 were from female F2; chicks A8 and A9 were sired by an extra-pair male. BM, CM, DM and EM were neighbouring males. Male BM sired chicks A3, A8 and A9; A1 and A2 were sired by male EM (confirmed by additional probes; Table 1).

METHODS. Full details of the cloning and fingerprinting methods are given elsewhere<sup>3,21,23</sup>. In order to assign the genetic paternity of each offspring, 5–10 µg *Mbol*-digested DNA with internal size marker (2.0 ng *Xho*I-digested  $\lambda$  DNA + 8.0 ng 1,012-bp ladder DNA (Gibco BRL)) was electrophoresed, blotted and probed sequentially with up to four <sup>32</sup>P-labelled single-locus minisatellite probes (cEscMS1, cEscMS2 and cEscMS7 cloned from reed bunting<sup>21</sup>; cGgaMS2 from chicken<sup>24</sup>), each in combination with <sup>32</sup>P-labelled 6.6-kb  $\lambda$ /*Hind*III fragment which hybridized to the 15.0-kb  $\lambda$ /*Xho*I internal marker band. The membranes were finally probed with <sup>32</sup>P-labelled internal marker DNA, and allele sizes estimated (to  $\pm 0.1$  kb for alleles <6.0 kb;  $\pm 0.2$  kb for alleles of 6.0–9.0 kb;  $\pm 0.5$  kb for alleles >9.0 kb) following alignment of the 15.0-kb marker bands on the test and marker autoradiographs. Females were run adjacent to their offspring in every case, allowing all maternal alleles to be identified; there was no intraspecific brood parasitism among the 216 chicks in 58 broods. Sets of paternally derived alleles were compared with the genotypes of all breeding males on the study area to identify fathers. The size 'bins' were very conservative to minimize the possibility of false exclusion of true fathers, especially when samples were run on different gels (Table 1). The analysis allowed the unambiguous assignment of 94% (203/216) of chicks to a single possible father; the remaining 6% could not be assigned to any territory-holding male in the study area.

\* Present address: Waltham Centre for Pet Nutrition, Melton Mowbray, Leicestershire LE14 4RT, UK.

consistency in the level of paternity achieved by males in successive broods (% EPP in nest 1 versus % EPP in nest 2: adjusted  $r^2 = 0.02$ , d.f. = 11,  $P = 0.39$ ). There was no tendency for second broods to contain a different proportion of extra-pair young than first broods; five second broods held fewer extra-pair offspring and seven held more (Fig. 2a). Neither was there a tendency for males to feed second broods at a lower rate; males fed at a lower rate in seven second broods and at a higher rate in five (Fig. 2a).

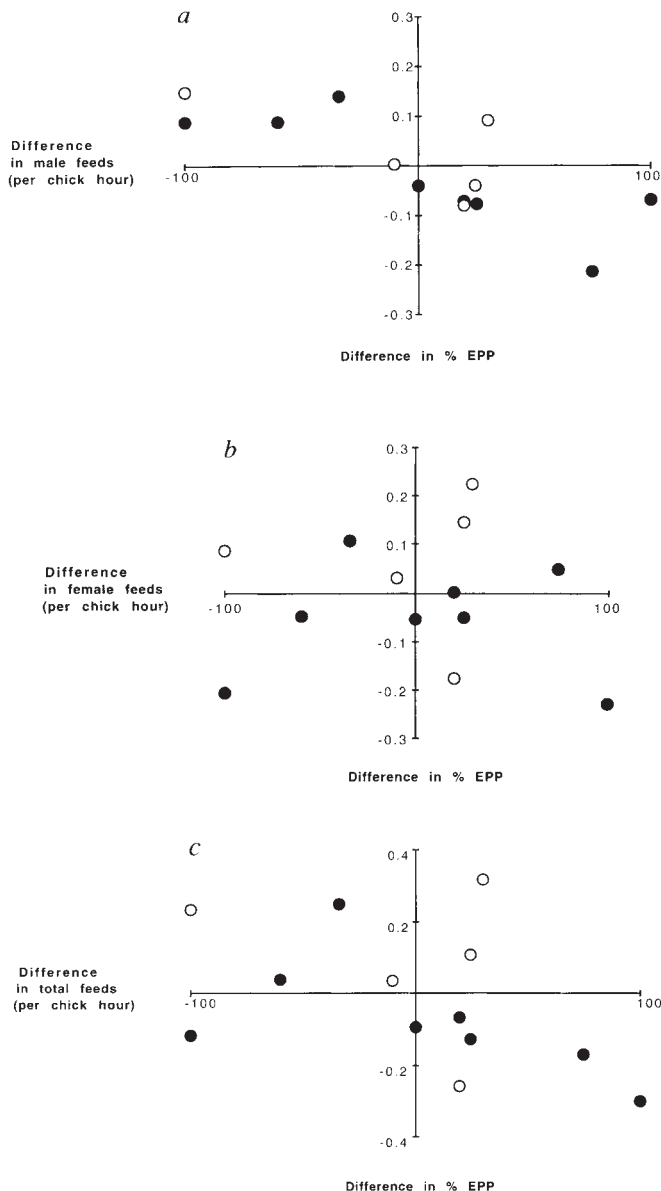


FIG. 2 Relationship between the difference in nestling provisioning rate and the difference in the proportion of extra-pair young in the first and second broods of double-brooded pairs. The vertical axis represents the difference in the mean residual feeds/hour per chick (after controlling for chick age using a least squares polynomial regression<sup>21</sup>) between first and second nests. Feeding rate per chick did not vary with brood size ( $r_s = 0.009$ ,  $N = 40$ ,  $P > 0.5$ ). Monogamous pairs' nests are indicated by filled circles, those of polygynous primary pairs by circles. The null hypothesis of no association between the change in feeding and change in paternity was tested using a two-tailed binomial test ( $N = 12$ , excluding one uninformative value). a, Male nestling provisioning rate and paternity,  $P = 0.0064$  (for monogamous nests only,  $P = 0.016$ ;  $N = 7$ ). b, Female nestling provisioning rate and paternity,  $P = 1.0$ . c, Total nestling provisioning rate (male + female) and paternity,  $P = 0.15$ .

TABLE 2 Variance of extra-pair paternity in male-assisted nests\*

| Brood size<br>(B) | N  | EPP<br>proportion | Variance              |                             | P†    |
|-------------------|----|-------------------|-----------------------|-----------------------------|-------|
|                   |    |                   | Observed<br>( $s^2$ ) | Expected‡<br>( $\sigma^2$ ) |       |
| 1                 | 1  | 1                 | —                     | —                           | —     |
| 2                 | 8  | 0.56              | 0.98                  | 0.50                        | 0.05  |
| 3                 | 11 | 0.48              | 1.47                  | 0.75                        | 0.05  |
| 4                 | 14 | 0.54              | 2.29                  | 1.00                        | 0.005 |
| 5                 | 13 | 0.51              | 1.27                  | 1.25                        | 0.5   |
| Combined§         |    |                   |                       |                             | 0.005 |

\* Includes nests of monogamous males and primary (first hatching) nests of polygynous males. These are the nests to which males direct their parental care. There was no significant difference in the frequency of extra-pair paternity between monogamous and primary nests<sup>21</sup>.

† Probabilities of  $\chi^2 = (N-1)s^2/\sigma^2$  found from the chi-square distribution with  $(N-1)$  d.f. (ref. 22).

‡ Binomial expected found as  $Bq(1-q)$ , where  $q$  = overall proportion of EPP (0.518) in primary and monogamous nests<sup>22</sup>.

§ Fisher's combined probability<sup>22</sup>.

A comparison between the difference in male feeding rate per chick and the difference in the proportion of extra-pair young between first and second broods revealed a highly significant relationship between the amount of care given by a male and his proportion of paternity (Fig. 2a). There was no relationship between female provisioning rate and paternity (Fig. 2b) and there was a nonsignificant trend towards broods containing a higher proportion of EPP receiving fewer feeds overall (Fig. 2c). This response to the degree of EPP provides compelling evidence that males adjust their parental effort in relation to their expected gametic contribution to the brood.

How do males assess their paternity success? Males feed broods composed entirely of extra-pair young, suggesting that males cannot recognize their own offspring. The most plausible explanation is that males somehow assess their paternity success during the female's fertile period. Evidence supporting this hypothesis has been found in the dunnoek, *Prunella modularis*<sup>18,19</sup> and swallow, *Hirundo rustica*<sup>13,16</sup>. In the dunnoek, males in polyandrous mating trios have been shown to use their share of matings with the female to determine their relative feeding rate<sup>17,18</sup>. In the monogamous swallow, males were found to allocate their investment in relation to both their absolute number of copulations and to their share of the total copulations engaged in by the female (that is, within-pair and extra-pair copulations).

Good genes hypotheses invoked to explain female copulation behaviour<sup>1</sup> require a heritable component of fitness which can be detected through a male's phenotype and predict that female genetic mate choice would be consistent among broods. However, the observed lack of consistency in the degree of EPP in successive nests does not concur with this prediction. The exceptionally high frequency of EPP also suggests that the low level of mate guarding observed in this population is not particularly effective<sup>20,21</sup>. Thus, it seems the best way for males to behave under such circumstances is to adapt their behaviour in relation to their perceived degree of cuckoldry. Recent theoretical models have predicted that such behaviour is adaptive when the expectation of paternity varies in successive nesting attempts<sup>8</sup>, as appears to be the case in the reed bunting. □

Received 23 March; accepted 6 September 1994.

1. Birkhead, T. R. & Møller, A. P. *Sperm Competition in Birds: Evolutionary Causes and Consequences* (Academic, London, 1992).
2. Wong, Z., Wilson, V., Patel, I., Povey, S. & Jeffreys, A. J. *Ann. hum. Genet.* **51**, 269–288 (1987).
3. Bruford, M. W., Hanotte, O., Brookfield, J. F. Y. & Burke, T. in *Molecular Genetic Analysis of Populations: A Practical Approach* (ed. Hoelzel, A. R.) 225–269 (IRL, Oxford, 1992).
4. Kempenaers, B. et al. *Nature* **357**, 494–496 (1992).
5. Smith, S. M. *Behaviour* **107**, 15–23 (1988).
6. Møller, A. P. *Behav. Ecol. Sociobiol.* **27**, 23–29 (1990).



7. Whittingham, L. A., Taylor, P. D. & Robertson, R. J. *Am. Nat.* **139**, 1115–1125 (1992).
8. Westneat, D. F. & Sherman, P. W. *Behav. Ecol.* **4**, 66–77 (1993).
9. Gavin, T. A. & Bollinger, E. K. *Auk* **102**, 550–555 (1985).
10. Frederick, P. *Behav. Ecol. Sociobiol.* **21**, 223–228 (1987).
11. Morton, E. S. *Behaviour* **101**, 211–224 (1987).
12. Westneat, D. F. *Auk* **105**, 149–160 (1988).
13. Møller, A. P. *Anim. Behav.* **36**, 996–1005 (1988).
14. Koenig, W. D. *Behav. Ecol. Sociobiol.* **1**, 55–61 (1990).
15. Morton, E. S., Forman, L. & Braun, M. *Auk* **107**, 275–283 (1990).
16. Møller, A. P. *Anim. Behav.* **42**, 261–268 (1991).
17. Burke, T., Davies, N. B., Bruford, M. W. & Hatchwell, B. J. *Nature* **338**, 249–251 (1989).
18. Davies, N. B., Hatchwell, B. J., Robson, T. & Burke, T. *Anim. Behav.* **43**, 729–745 (1992).
19. Davies, N. B. *Dunnock Behaviour and Social Evolution* (Oxford University Press, Oxford, 1992).
20. O'Malley, S. L. C. thesis, Univ. Leicester (1993).
21. Dixon, A. thesis, Univ. Leicester (1993).
22. Sokal, R. R. & Rohlf, F. J. *Biometry* (Freeman, San Francisco, 1981).
23. Hanotte, O., Burke, T., Armour, J. A. L. & Jeffreys, A. J. *Genomics* **9**, 587–597 (1991).
24. Bruford, M. W., Hanotte, O. & Burke, T. *Anim. Genet.* (in the press).

ACKNOWLEDGEMENTS. We thank the Leicestershire and Rutland Naturalists Trust and T. Appleton for permission to work at the field site; M. Lester for assistance in the field; O. Hanotte and D. Thomaz for advice in the laboratory; and K. Lessels for comments. This work was supported by the NERC, the SERC and the Royal Society.

## Why two eyes are better than one for judgements of heading

A. V. van den Berg & E. Brenner

Physiology I, Medical Faculty, Erasmus University Rotterdam,  
PO Box 1738, 3000 DR Rotterdam, The Netherlands

ARE two eyes needed for judging direction of self-motion? Traditional analyses stress that the pattern of optic flow in one eye is sufficient<sup>1–5</sup>. The main difficulty is how to deal with the eye or head rotation. Extraretinal signals help<sup>6–8</sup>, but humans can also discount the effect of rotation purely on the basis of monocular flow<sup>6,7,9–12</sup> provided the scene contains depth<sup>6,9,10</sup>. Depth differences give rise to changing binocular disparities when the observer moves. These disparities are ignored in monocular theories of judgements of heading. Using computer generated displays, we investigated whether stereoscopic presentation improves heading judgements for conditions that pose problems to the monocular observer. We found that adding disparities to simulated ego-motion through a cloud of dots made heading judgements up to four times more tolerant to motion noise. The same improvement was found when the disparities specify the initial distances throughout the motion sequence. We conclude that binocular disparities improve judgements of heading by imposing a depth order on the elements of the scene, not because they provide additional information on the elements' motion in depth.

When a driver fixates a mountain ridge in the distance, his direction of gaze is practically stationary, and the retina receives a motion pattern that radiates outward from his direction of heading. In contrast, when rotating his eye and his head so as to fixate a road sign, he will null the sign's motion on the fovea. In this case, the retinal motion pattern (retinal flow) radiates outward from the fixation point rather than from the destination point. How can humans disregard the rotational component, which complicates the judgement of heading? Normally, visual and extraretinal signals that accompany the self rotation work in concert to discount the rotation<sup>6–8</sup>. Nevertheless, when one presents the retinal motion of a rotating and translating observer to a stationary eye, heading is often perceived accurately<sup>6,7,9–12</sup>. Under such conditions, monocular heading judgements are sensitive to the layout of the environment. They are accurate in the presence of noise<sup>7,12</sup> or fast eye rotations<sup>11</sup> when motion across the ground plane is simulated, but not for motion through a cloud of dots<sup>7,8</sup>. Depth cues (perspective, texture gradients and height in the display) help to derive the heading from the retinal flow in the case of the ground plane<sup>12</sup>. Hence, we surmised that

adding stereoscopic information to the flow would improve the performance for motion through the cloud.

We investigated the sensitivity of heading judgements to noise. In the first experiment we used presentations with and without stereoscopic information, both for the ground plane and for a cloud of dots. Horizontal simulated self-motion was always presented to both eyes, but in the synoptic (monocular information) case the eyes received identical images. We simulated the version and, in the stereoscopic condition, the vergence eye movements that were required to fixate a point in the environment. At the end of the motion sequence, the subject used a pointer to indicate the perceived direction of heading (Fig. 1).

For stereo presentations we found similar performance for simulated motion across the plane and for motion through the cloud (see example in Fig. 2). For motion across the ground plane, we found little difference between stereoscopic and synoptic presentation. In both cases, pointing was accurate and precise when the speed of the local motion vectors in the flow was four times larger than the speed of the local noise (signal-to-noise ratio = 4, see Fig. 1 legend). For lower SNR, precision decreased and subjects showed an increasing tendency to point towards the fixation point. Little correlation remained between the pointing responses and the simulated heading direction when the noise exceeded the signal (SNR < 1.0).

For motion through a cloud of dots, we found a clear difference between stereoscopic and synoptic presentations (circles in Fig. 3). For stereoscopic presentation heading was perceived

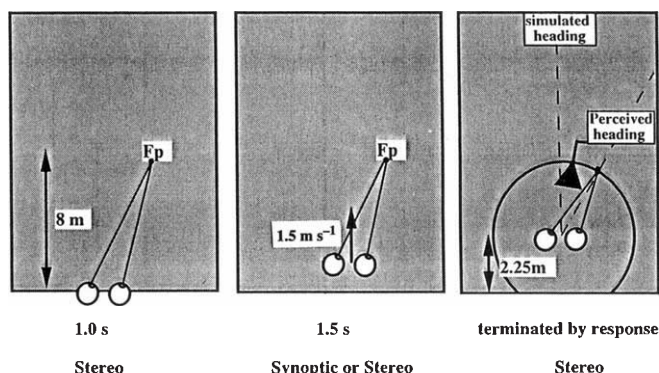


FIG. 1 The sequence of events during a trial. The scene contained about 256 white dots randomly distributed in a cloud or on the ground plane. The simulated depth range was from 1 to 20 m. Other dimensions were determined by the screen size (60° horizontally × 50° vertically). Subjects fixated a red point, that was part of the scene, at variable eccentricity and initially at 8 m distance. The simulated eye height above the ground plane was 0.65 m. To aid fixation, the first frame was shown stereoscopically for one second. Subsequently, forward motion was simulated with a speed of 1.5 m s<sup>-1</sup>. During this period presentation was either synoptic or stereoscopic. The synoptic motion sequence corresponded to the motion pattern that would be received by a point at the bridge of the nose. Dot lifetime was limited to 160 ms to rule out the use of cues related to the trajectories of individual dots. Each dot's motion was perturbed with randomly directed noise. The magnitude of the noise component was proportional to the local flow velocity (SNR =  $v_{\text{flow}}/v_{\text{noise}}$ ). The eye rotations required to fixate the red point (Fp) were simulated. Thus, the images of the red point for the two eyes were stationary on the screen. This imposed a fixed eye vergence that corresponded to a distance halfway between the initial and the final simulated positions of the red point. After 1.5 s the motion stopped. The scene was shown stereoscopically, with a triangular pointer, which the subject turned about a circle concentric with his feet so as to indicate the perceived direction of heading. A button press terminated the presentation. Subjects were told that the displays mimicked the view one would receive when looking at a road sign while driving a car. They were asked to indicate the heading direction of the car. All three subjects were given feedback on their performance during 10–50 training trials, but not during testing.

down to SNR = 1. For synoptic presentation, subjects could tolerate less noise and perceived heading down to SNR = 2 or SNR = 4. Note that as more noise was introduced (lower SNR), the precision decreased and subjects' responses became more biased towards the fixation point (Fig. 3).

In the second experiment we investigated whether changing disparity was essential for the improved performance in the cloud. The motion sequence was identical in the two eyes, but each dot of the cloud was given a fixed disparity that corresponded to the dot's simulated three-dimensional position in the first frame. Performance for this 'static-stereo' presentation was very similar to that for the full-stereo condition (unfilled symbols in Fig. 3). Thus, the depth order that static disparities impose on the dots in the cloud is sufficient to enhance performance to the level attained for the ground plane.

We have shown that stereoscopic depth is beneficial to the perception of heading. This is the case, not because it provides a direct cue to the motion in depth of the rigid environment relative to the observer, but because it provides a depth order, as could occlusion or texture gradients. The relative magnitude of the translatory and the rotatory contributions to the flow-field changes with the distance. We think that the independent information on the depth order allows the brain to exploit this property of the flow-field. Specifically, the most distant points are most reliable for estimation of the self-rotation, because the translatory part of the flow is inversely proportional to the distance. Conversely, for relatively small rotation rates the flow of nearby points is dominated by the translatory part. Such relations may be used to constrain the set of possible heading directions and ego-rotations that are consistent with the

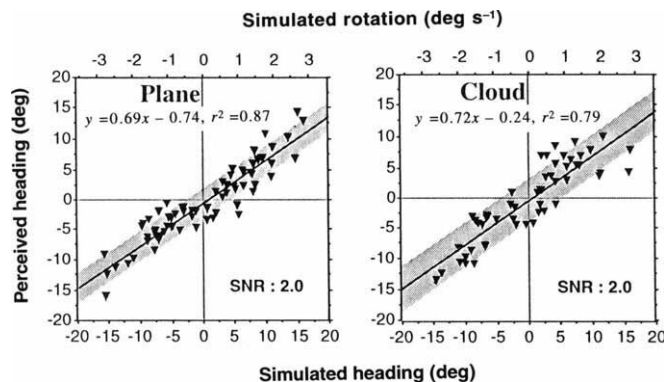


FIG. 2 Example of pointing responses for simulated motion across the plane and for motion through the cloud. Each point indicates the response in a single trial. Heading is expressed as an angle relative to the fixation direction at the end of the presentation. Perceived and simulated heading directions are linearly related. If the correlation exceeds a criterion level (0.5), we characterize the constant and the variable error of the subject's pointing response by the slope of the regression line and the s.d. of the perceived heading relative to the predicted heading (s.d.( $\epsilon$ )) using the regression line. Perfect pointing would correspond to a slope of 1.0 and negligible variation of pointing (s.d.( $\epsilon$ ) = 0). This subject's (J.K.) perceived heading was biased towards the fixation point by about 30%, irrespective of the layout of the points (the slope of the regression line is about 0.7 in both cases). His variable part of the pointing error was also very similar (cloud: 2.9°; plane: 2.2°). The shaded area in each panel indicates a range of 2 s.d.( $\epsilon$ ) about the best fitting regression line. The upper axis indicates the average simulated horizontal rotation rate.

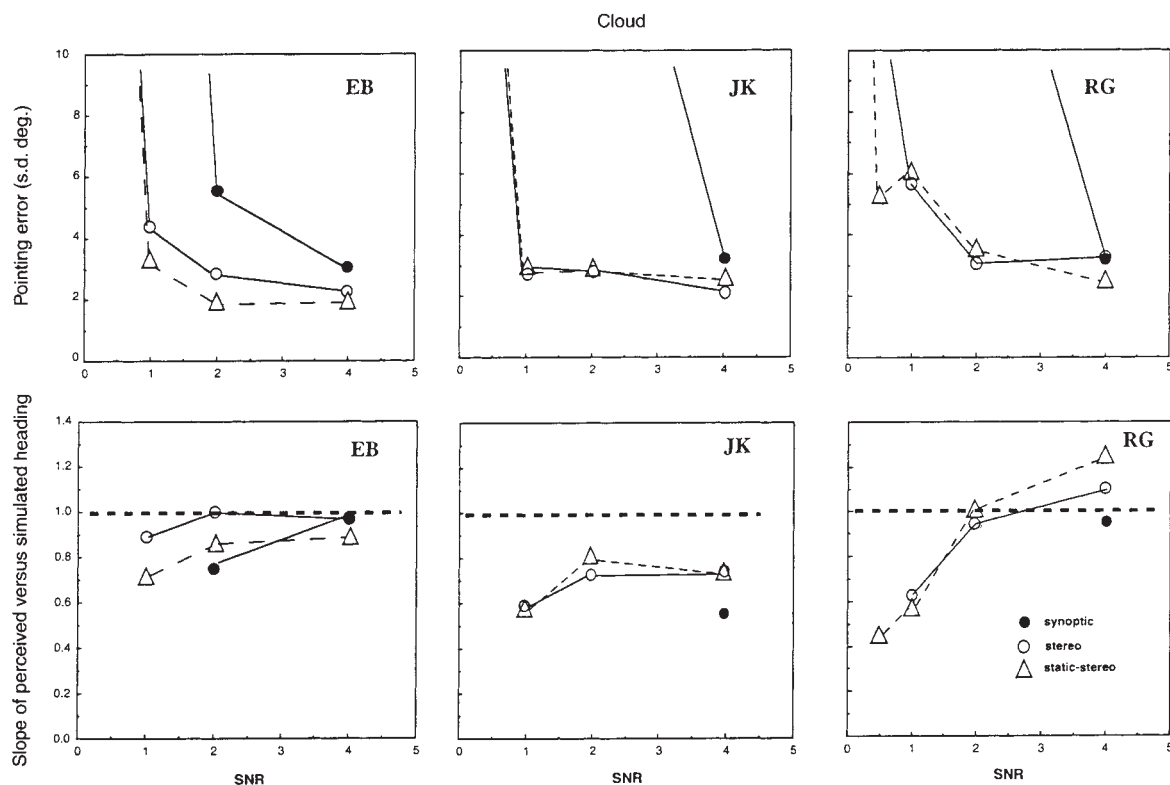


FIG. 3 How the pointing responses depend on the SNR and the type of information presented in the displays. Motion through the cloud was simulated. The subjects' variable error (s.d.( $\epsilon$ )) is indicated in the upper panels. The steep upward lines indicate that at the next lower SNR

level, correlation between pointing responses and simulated heading was less than 0.5. The lower panels show the slopes of the perceived versus the simulated heading. Values lower than one indicate a bias towards the fixation point.



observed flow-field, resulting in reduced scatter in the perceived heading. Without static depth information, visual heading judgements are more vulnerable to noise and the confounding effects of eye and head rotation. □

Received 30 March; accepted 26 August 1994.

1. Gibson, J. J. *Perception of the Visual World* (Houghton Mifflin, Boston, 1950).
2. Longuet-Higgins, H. C. & Prazdny, K. *Proc. R. Soc. B* **208**, 385–397 (1980).
3. Rieger, J. H. & Lawton, D. T. *J. opt. Soc. Am.* **A2**, 354–359 (1985).
4. Koenderink, J. J. & Van Doorn, J. A. *Biol. Cybern.* **56**, 247–254 (1987).
5. Heeger, D. J. & Jepson, A. *Neural Comp.* **2**, 129–137 (1990).
6. Warren, W. H. & Hannon, D. J. *Nature* **336**, 162–163 (1988).
7. van den Berg, A. V. *Vision Res.* **32**, 1285–1296 (1992).
8. Roeyden, C. S., Banks, M. S. & Crowell, J. A. *Nature* **360**, 583–585 (1992).
9. Warren, W. H. & Hannon, D. J. *J. opt. Soc. Am.* **A7**, 160–169 (1990).
10. Rieger, J. H. & Toet, L. *Biol. Cybern.* **52**, 377–381 (1985).
11. van den Berg, A. V. *Nature* **365**, 497–498 (1993).
12. van den Berg, A. V. & Brenner, E. *Vision Res.* **34**, 2153–2167 (1994).

## β-Adrenergic activation and memory for emotional events

Larry Cahill\*, Bruce Prins†, Michael Weber†‡ & James L. McGaugh\*

\* Center for the Neurobiology of Learning and Memory, and Department of Psychobiology, University of California, Irvine, California 92717-3800, USA

† Hypertension Center, Long Beach Veteran's Affairs Medical Center, Long Beach, California 90822, USA

‡ Department of Medicine, University of California, Irvine, California 92717-4075, USA

**SUBSTANTIAL evidence from animal studies suggests that enhanced memory associated with emotional arousal results from an activation of β-adrenergic stress hormone systems during and after an emotional experience<sup>1–3</sup>. To examine this implication in human subjects, we investigated the effect of the β-adrenergic receptor antagonist propranolol hydrochloride on long-term memory for an emotionally arousing short story, or a closely matched but more emotionally neutral story. We report here that propranolol significantly impaired memory of the emotionally arousing story but did not affect memory of the emotionally neutral story. The impairing effect of propranolol on memory of the emotional story was not due either to reduced emotional responsiveness or to nonspecific sedative or attentional effects. The results support the hypothesis that enhanced memory associated with emotional experiences involves activation of the β-adrenergic system.**

Subjects received either propranolol or a placebo 1 h before viewing a series of slides accompanied by an emotional or neutral narrative, and were tested for memory of the story one week later (see Fig. 1 for methods). The stories were those used in an earlier study demonstrating an enhancing effect of emotional arousal on memory (L.C. and J.L. McG., manuscript in preparation), and were developed from earlier work by other investigators demonstrating enhancing effects of emotional arousal on memory<sup>4</sup> (Box 1). If the enhanced memory for the emotional story involved activation of β-adrenergic receptors (either centrally or peripherally), then blockade of those receptors should impair memory for the emotional story, while leaving memory for the neutral story relatively unaffected. Our results confirm this prediction.

The memory test results revealed significant and selective effects of propranolol on memory of the emotional story. Focusing first on the free recall results, we examined the mean number of slides recalled (out of 12 possible). The placebo subjects who viewed the emotional story recalled significantly more slides ( $6.0 \pm 0.6$ ) than did propranolol subjects ( $4.09 \pm 0.55$ ) ( $t(17) = 2.33$ ,  $P < 0.05$ ). In contrast, the placebo and propranolol groups

who viewed the neutral story did not differ in the number of slides recalled. Similar results were obtained in an 80-item multiple-choice recognition memory test (which assessed memory for both visual and narrative story elements). Placebo subjects who viewed the emotional story answered significantly more questions correctly ( $48.9 \pm 1.47$ ) than did the propranolol subjects ( $42.4 \pm 1.72$ ) ( $t(17) = 2.73$ ,  $P < 0.02$ ). The placebo and propranolol groups who viewed the neutral story did not differ in number of questions correctly answered.

The enhancing effects of emotional activation on recognition memory and the impairing effects of propranolol on emotionally enhanced memory were obtained primarily in story phase 2, the phase in which the emotional elements were introduced (Fig. 1). The placebo subjects displayed superior memory for those story elements associated with emotional arousal, whereas the propranolol subjects did not. A 2-factor ANOVA for the arousal story results with repeated measures on the story phase revealed significant effects of the drug treatment ( $P < 0.05$ ) and story phase ( $P < 0.01$ ). Subjects in the placebo, arousal story condition answered significantly more phase 2 questions correctly than either phase 1 ( $P < 0.01$ ) or phase 3 ( $P < 0.05$ ) questions. Furthermore, and most importantly, the retention performance of the placebo group was significantly better than that of the propranolol group for questions pertaining to phase 2 of the arousal story ( $P < 0.02$ ). In contrast, for subjects in the propranolol/arousal story condition, as well as subjects in the placebo and propranolol groups given the neutral story, the

### BOX 1 Narratives accompanying slide presentation

| Slide | Neutral version                                                                                           | Arousal version                                                                                                 |
|-------|-----------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------|
| 1.    | A mother and her son are leaving home in the morning.                                                     | A mother and her son are leaving home in the morning.                                                           |
| 2.    | She is taking him to visit his father's workplace.                                                        | She is taking him to visit his father's workplace.                                                              |
| 3.    | The father is a laboratory technician at Victory Memorial Hospital.                                       | The father is a laboratory technician at Victory Memorial Hospital.                                             |
| 4.    | They check before crossing a busy road.                                                                   | They check before crossing a busy road.                                                                         |
| 5.    | While walking along, the boy sees some wrecked cars in a junk yard, which he finds interesting.           | While crossing the road, the boy is caught in a terrible accident, which critically injures him.                |
| 6.    | At the hospital, the staff are preparing for a practice disaster drill, which the boy will watch.         | At the hospital, the staff prepare the emergency room, to which the boy is rushed.                              |
| 7.    | An image from a brain scan machine used in the drill attracts the boy's interest.                         | An image from a brain scan machine used in a trauma situation shows severe bleeding in the boy's brain.         |
| 8.    | All morning long, a surgical team practised the disaster drill procedures.                                | All morning long, a surgical team struggled to save the boy's life.                                             |
| 9.    | Make-up artists were able to create realistic-looking injuries on actors for the drill.                   | Specialized surgeons were able to re-attach the boy's severed feet.                                             |
| 10.   | After the drill, while the father watched the boy, the mother left to phone her other child's pre-school. | After the surgery, while the father stayed with the boy, the mother left to phone her other child's pre-school. |
| 11.   | Running a little late, she phones the pre-school to tell them she will soon pick up her child.            | Feeling distraught, she phones the pre-school to tell them she will soon pick up her child.                     |
| 12.   | Heading to pick up her child, she hails a taxi at the number nine bus stop.                               | Heading to pick up her child, she hails a taxi at the number nine bus stop.                                     |

recognition test scores for phase 2 did not differ from those of phases 1 or 3.

The subjects' heart rate and blood pressure were measured immediately before administration of the placebo or propranolol and again before presentation of the story. As expected (Table 1), propranolol significantly decreased heart rate and reduced blood pressure.  $\beta$ -Blockers have been reported to induce side effects, including sedation and difficulty in focusing attention in some patients<sup>3,6</sup>. If propranolol treatment impaired memory in the present study simply by increasing sedation or impairing attention, then memory for both the neutral and arousal stories should have been affected. That was clearly not the case: propranolol treatment selectively impaired memory for the more

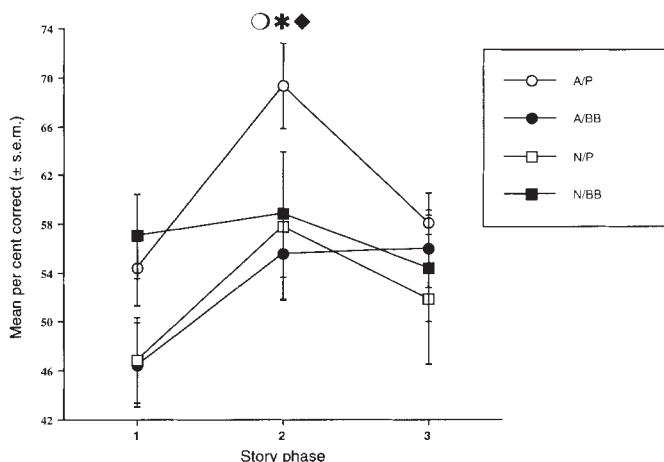


FIG. 1 Results of the recognition (multiple-choice) memory test for the three phases of the arousal and neutral stories. A/P, arousal story/placebo treatment; A/BB, arousal story/ $\beta$ -blocker treatment; N/P, neutral story/placebo treatment; N/BB, neutral story/ $\beta$ -blocker treatment. ANOVA indicated significant effects of drug treatment ( $F(1, 17)=6.62$ ,  $P<0.05$ ) and story phase ( $F(2, 34)=6.84$ ,  $P<0.01$ ). Placebo subjects given the arousal story had enhanced memory for phase 2 (arousing phase) of the story:  $\circ$ ,  $<0.01$  compared to phase 1 arousal story/placebo scores ( $t(14)=3.18$ ) and  $\diamond$ ,  $<0.05$  compared to phase 3 arousal story/placebo scores ( $t(11)=2.58$ ). Propranolol blocked the enhancing effect of arousal on memory:  $\blacklozenge$ ,  $P<0.02$  compared to arousal/ $\beta$ -blocker group story phase two ( $t(17)=2.58$ ). Results are expressed as per cent of questions answered correctly because of different numbers of questions given for each phase. Subjects (19 females, 17 males, mean age ( $\pm$ s.e.m.)=27.4( $\pm$ 4.6) years) received a placebo or propranolol hydrochloride tablet (Inderal, 40 mg; Wyeth-Ayerst Laboratories) 1 h before viewing either an emotionally arousing story, or a closely matched but more emotionally neutral story (Box 1). There were 8–11 subjects per group (the differences in group sizes were due primarily to failure of some subjects to appear for the memory test session). The stories were presented as a brief (about 4 min) narrated slide show. The one-sentence narrations were identical for the first 4 (of 12) slides (referred to as phase 1), nearly identical for the last three slides (phase 3), and differed primarily in the middle five slides (phase 2), during which the emotional events were introduced into the arousing story. Subjects were blind to the drug treatment received, and were tested individually. They were connected to heart rate and blood pressure monitors while viewing the story, and were told that the study concerned physiological responses to different types of stimuli. Testing was conducted between 9:00 am and 1:00 pm. One week later the subjects were given surprise memory tests. A free recall test, in which subjects were asked to recall as much as possible of the story, was followed by a 4-choice multiple-choice recognition test consisting of 5–8 questions per slide pertaining to visual and narrative elements for each story slide. For each question, one correct and three plausible (but incorrect) alternatives were provided. The free recall responses were tape-recorded for subsequent scoring. In the free recall test, subjects were credited with recall of a slide if they described some story information that could only have been known from viewing that slide. The entire study (including scoring of the recall tapes) was conducted with the experimenter blind to the drug condition.

TABLE 1 Propranolol effects on heart rate and blood pressure

|                        | $\Delta$ Heart rate<br>(beats per min,<br>mean $\pm$ s.e.m.) | $\Delta$ Mean blood<br>pressure (mm Hg,<br>mean $\pm$ s.e.m.) |
|------------------------|--------------------------------------------------------------|---------------------------------------------------------------|
| Placebo ( $n=15$ )     | 1.7 ( $\pm$ 2.41)                                            | 0.11 ( $\pm$ 2.4)                                             |
| Propranolol ( $n=20$ ) | -7.79 ( $\pm$ 1.68)*                                         | -8.74 ( $\pm$ 0.94)*                                          |

The scores are based on differences between measures taken immediately before drug or placebo administration and immediately before presentation of the slides and narration. Propranolol produced the expected decreases in baseline heart rate and blood pressure (for both measures  $P<0.0005$  from placebo group ( $t$ -test)). Propranolol hydrochloride blocks both  $\beta_1$  and  $\beta_2$  adrenergic receptors and readily crosses the blood-brain barrier. It is widely used for the treatment of certain cardiac disorders such as angina pectoris and high blood pressure.<sup>15</sup> Propranolol was used in this study because it impairs  $\beta$ -adrenergic transmission both centrally and peripherally and because it attenuates memory in animal experiments using aversively motivated training tasks<sup>16–18</sup>.

emotional story. Furthermore, as propranolol did not block subjects' subjective emotional reactions to the story assessed immediately after story viewing (Fig. 2), it is clear that the effects of propranolol on memory reported here are not easily attributable to reduced subjective emotional responsiveness.

The results support the hypothesis that memory storage is modulated by  $\beta$ -adrenergic systems. In particular, they suggest that the enhanced memory for events associated with emotional arousal involves activation of  $\beta$ -adrenergic receptors and that  $\beta$ -adrenergic activation is not essential for storage of emotionally neutral information. It is well established that highly emotional experiences can activate the sympathetic nervous system as evidenced, for example, by increases in plasma adrenaline and/or noradrenaline<sup>7,8</sup>. There is also extensive evidence indicating that, in animals, administration of adrenergic and other stress-related hormones and drugs can modulate (either enhance or impair) memory in a wide variety of learning situations<sup>9</sup>. Such findings suggest that emotional experiences activate a memory-modulating system involving the release of both peripheral adrenergic hormones and brain noradrenaline<sup>3,10</sup>. Although there is extensive evidence that high doses of adrenaline can impair memory storage<sup>1,9</sup>, it is not known whether endogenous adrenaline

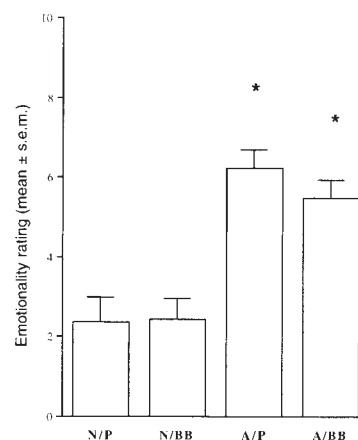


FIG. 2 Mean ( $\pm$ s.e.m.) ratings of emotional reaction to the story viewed in each experimental group. N/P, neutral story, placebo; N/BB, neutral story,  $\beta$ -blocker; A/P, arousal story, placebo; A/BB, arousal story,  $\beta$ -blocker. \*,  $P<0.01$  from both N/P and N/BB groups ( $t$ -tests). Immediately after viewing the story, subjects rated how emotional they found the story to be on a scale of 0 to 10 (with 0 as 'not emotional at all' and 10 as 'highly emotional') by marking a scale at the appropriate point.



released by high levels of emotional arousal can produce memory impairment. Findings of animal experiments indicate that memory storage is influenced by drugs affecting many neuromodulatory systems, including  $\alpha$ -adrenergic systems<sup>1,3,9,11</sup>. Studies to date have not investigated the selective involvement of other systems in regulating emotionally influenced memory in human subject.

Although clinical reports suggest that  $\beta$ -blockers may induce memory loss in some patients<sup>12,13</sup>, controlled laboratory investigations have so far generally failed to find consistent effects of  $\beta$ -blockers on memory<sup>6,14</sup>. However, previous studies generally investigated memory for relatively unemotional events and typically studied only short-term memory. Future research using  $\beta$ -blockers with different affinities for  $\beta_1$  and  $\beta_2$  receptors and differential effectiveness in passing the blood-brain barrier should clarify the role of the  $\beta$ -adrenergic system in regulating long-term memory for emotional experiences.  $\square$

Received 14 June; accepted 13 September 1994.

- McGaugh, J. L. *A. Rev. Neurosci.* **12**, 255–287 (1989).
- McGaugh, J. L. *et al. Neurosci. Biobehav. Rev.* **14**, 425–431 (1990).
- McGaugh, J. L. in *Handbook of Emotion and Memory: Research and Theory* (ed. Christianson, S. A.) 245–268 (Erlbaum, Hillsdale, 1992).
- Heuer, F. & Reisberg, D. *Memory Cognition* **18**, 496–506 (1990).
- Turkkan, J. *Neurosci. Biobehav. Rev.* **12**, 111–122 (1988).
- Dimsdale, J., Newton, R. & Joist, T. *Arch. int. Med.* **149**, 514–525 (1989).
- Frankenhaeuser, M. in *Emotions: Their Parameters and Measurement* (ed. Levi, L.) 209–234 (Raven, New York, 1975).
- McCarty, R. & Gold, P. E. *Horm. Behav.* **15**, 168–182 (1981).
- McGaugh, J. L. & Gold, P. E. in *Psychoendocrinology* (eds Brush, R. B. & Levine, S.) 305–339 (Academic, New York, 1989).
- McGaugh, J. L. *A. Rev. Psychol.* **34**, 297–323 (1983).
- Arnst, A. F. T. & Goldman-Rakic, P. S. in *Memory Dysfunctions: An Integration of Animal and Human Research from Clinical and Preclinical Perspectives* (eds Oton, D. S., Gamzu, E. & Corkin, S.) 218–241 (N.Y. Acad. Sci., 1985).
- Dennis, K., Froman, D., Morrison, A., Holmes, K. & Howes, D. *Heart Lung* **20**, 459–463 (1991).
- Solomon, S. *et al. Arch. gen. Psychol.* **40**, 1109–1112 (1983).
- McAinsh, J. & Cruikshank, J. M. *Pharmac. exp. Ther.* **46**, 163–197 (1990).
- Nies, A. S. in *Cardiovascular Drug Therapy* (ed. Messerli, F. H.) 420–441 (Saunders, Philadelphia, 1990).
- Gold, P., Vogt, J. & Hall, J. *Behav. Neural Biol.* **46**, 145–155 (1986).
- McGaugh, J. L. in *Memory Dysfunctions: An Integration of Animal and Human Research from Preclinical and Clinical Perspectives* (eds Oton, D. S., Gamzu, E. & Corkin, S.) 150–161 (N.Y. Acad. Sci., New York, 1985).
- Gallagher, M. *et al. in The Amygdaloid Complex* (ed. Ben-Ari, Y.) 343–354 (Elsevier/North Holland, Amsterdam, 1981).

## Changes in reliability of synaptic function as a mechanism for plasticity

Charles F. Stevens & Yanyan Wang

Laboratory for Molecular Neurobiology,  
Howard Hughes Medical Institute, Salk Institute,  
10010 North Torrey Pines Road, La Jolla, California 92110, USA

**SYNAPTIC transmission in the hippocampus is rather unreliable, with many presynaptic action potentials failing to release neurotransmitter<sup>1–4</sup>. How is this unreliability affected by the alterations in synaptic strength seen in long-term potentiation (LTP)<sup>5</sup> and long-term depression<sup>6,7</sup> (LTD)? We find that LTP increases synaptic reliability, and LTD decreases it, both without a change in the size of those postsynaptic currents that do occur. Thus LTD is a functional inverse of LTP.**

We have used 'minimal stimulation', a method for activating only one or a few synapses<sup>1</sup>: the stimulus intensity is reduced to a level that produces a postsynaptic response of uniform latency and shape in a target neuron with a probability of less than about 0.5. This method permits us to separate two components of synaptic strength: reliability, the fraction of stimuli that pro-

duce a postsynaptic current, and potency, the average peak size of the postsynaptic current when one does occur.

Whole-cell recording<sup>8,9</sup> provides a signal-to-noise ratio that is adequate for clearly distinguishing between a release failure and a success in almost all instances. This is illustrated in Fig. 1, which reveals a clear separation between transmitter release and no release.

Data for typical LTP experiments are illustrated in Figs 1*a* and 2*a*. Separate averages were made (before and after the induction of LTP) for all responses, all stimulation trials for which a postsynaptic current was detected, and all records that represented failures in synaptic transmission (Fig. 2*a*). Further, non-overlapping groups of about 25 successive traces were averaged throughout the experiment, and failure probabilities were estimated for the same groups of traces; these group averages reveal stable LTP following the tetanic stimulus (Fig. 2*b*), and a sustained post-tetanic decrease in the failure rate (Fig. 2*c*). The group average potency is roughly constant throughout the experiment (Fig. 2*d*), and the histograms of excitatory postsynaptic potential (e.p.s.c.) peak amplitude for the control and LTP periods (Fig. 2*e*) are not significantly different (Kolmogorov-Smirnov test,  $P > 0.1$ ).

In this instance, stable LTP is represented as an increase in the average response, a decrease in the synaptic failure rate, and no appreciable change in the potency. As with the LTP experiment, LTD is represented by a decrease in the average response size (Fig. 3*a, b*), an increase in the failure rate (Fig. 3*c*), and no appreciable change in synaptic potency (Fig. 3*a, d, e*).

Although we find, as reported previously<sup>10</sup>, that LTP is difficult to induce after about 20 min of recording in the whole-cell mode, LTD seems not to 'wash out'. That is, we could produce LTD as late as 80 min after the start of whole-cell recording, the longest time tested. In six experiments, LTD reversed previously established LTP (data not shown).

To compare data from different experiments, we note that the average peak e.p.s.c.,  $r_j$ , is (by definition) the product of the probability  $w_j$  that a release will occur and the average size  $a_j$  of a response (when release does occur):  $r_j = w_j a_j$ ; the subscript  $j$  is 0 before the induction of LTP or LTD and 1 after. The probability of a failure  $f_j = (1 - w_j)$  is determined directly from experiment as is the average potency  $a_j$ . Synaptic plasticity  $S$  is defined as the ratio  $S = r_1/r_0$ . If plasticity were only a change in synaptic reliability (with no change in potency), then  $a_1/a_0$  would equal 1, and  $S$  would be just the ratio of success probabilities:  $S = w_1/w_0$ , if  $a_1 = a_0$ . A plot of  $S$  versus the ratio  $w_1/w_0$  would then fall on a straight line with slope 1, and a plot of the ratio  $a_1/a_0$  would be constant ( $=1$ ). Our experiments provide the data required for testing the extent to which plasticity is determined by changes in reliability and changes in potency as we estimate  $a_1/a_0$  and  $w_1/w_0$  directly.

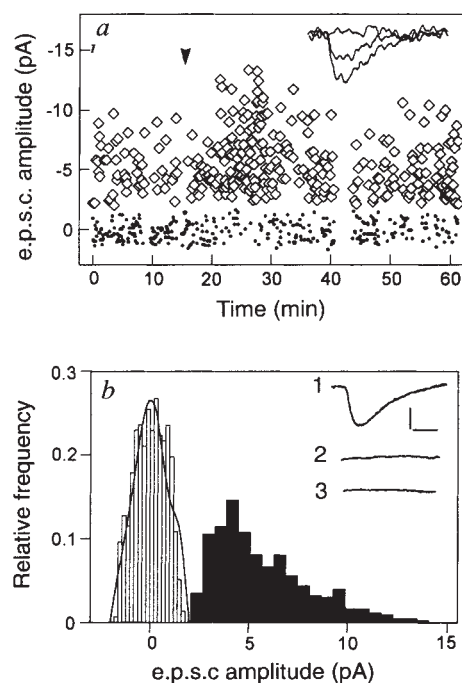
The unfilled circles in Fig. 4*a* and *b* represent data from 26 experiments like the ones illustrated in Figs 1, 2 and 3. Clearly, the ratio  $w_1/w_0$  falls along the diagonal in Fig. 4*a*, and the ratio  $a_1/a_0$  is roughly constant in Fig. 4*b*. For these experiments, then, we conclude that the mechanism of LTP and LTD is a change in synaptic reliability with no appreciable change in potency.

The experimental situation here is a little different from previous investigations of LTP that activated a population of synapses so that all or most failures are obscured. Because we have activated fewer synapses than was usual for earlier studies, we felt that we must demonstrate that the plasticity examined here exhibits the same properties as standard LTP.

LTP is sometimes elicited by the 'pairing' stimulation model instead of by tetanic stimulation. In these 'pairing' experiments, low-frequency stimulation is paired with a postsynaptic depolarization imposed on the cell<sup>11</sup>. Data from five 'pairing' experiments, plotted in Fig. 4*c* and *d* as unfilled triangles, conform to the behaviour of synapses potentiated with tetanic stimulation.

LTP is blocked by the presence of the *N*-methyl-D-aspartate (NMDA) receptor antagonist AP-5<sup>12</sup>, and by hyperpolarization

FIG. 1 *a*, Amplitude of synaptic currents as a function of time during the experiment to estimate the uncertainty in classifying traces as 'release' or 'no release'. Each point was calculated by averaging the current over a fixed window 2 ms long centred at the time that corresponded to the peak current. Diamonds are from traces on which release was apparent, and dots from traces without detectable release (currents averaged over the same 2 ms window). Tetanic stimulation was given at the time indicated by the arrow, and produced LTP of 170%. Three typical traces of current as a function of time, one without release and two with release, are superimposed in the inset. The calibration bars in *b* inset apply to these traces. *b*, Amplitude histogram for the data presented in *a*. The filled histogram was derived from all of the traces, represented by diamonds in which the experimenter could detect a synaptic current, and the open bars from all of the release failures (dots in *a*). The smooth histogram was derived from traces measured in the same way (and from the same experiment) except no stimulus was present to evoke synaptic current. The 'failures' histogram (unfilled bars) and the 'no stimulus' histogram (smooth line) are not significantly different (Kolmogorov-Smirnov test,  $P > 0.1$ ). The inset presents average current as a function of time from (1) all data in the filled histogram, (2) all data in the unfilled bar histogram, and (3) all data from the line histogram. We estimate from the overlaps of the 'releases' and 'failures' histograms, and from uncertainties in making decisions with visual classification and in this and other experiments, that less than about 5% of the traces would have been misclassified. In most experiments, the classification of traces as 'release' or 'no release' was done visually by the experimenter so that every trace was scrutinized. Most classification was done blind in the sense that the experimenter did not know if the records came from an LTP or an LTD experiment or if the traces being classified were obtained before or after the conditioning event (tetanus, pairing or low-frequency stimulation). In about 20% of the instances a second experimenter classified the traces independently; classifications by two independent experiments agree to within a few per cent. The calibration bars apply to the three inset traces in *a* and *b*. Vertical bar, 3 pA; horizontal bar, 5 ms. METHODS. Slices were prepared as in refs 16 and 17. Transverse hippocampal slices (~400  $\mu\text{m}$  thick) were obtained from rats (P14-P26, mostly P15-P24, Harlan Sprague-Dawley). The temperature in the recording chamber was  $31.5 \pm 0.5^\circ\text{C}$ . The extracellular solution is composed of (in mM): NaCl, 120; KCl, 3.5;  $\text{NaH}_2\text{PO}_4$ , 1.25;  $\text{NaHCO}_3$ , 26;  $\text{MgCl}_2$ , 1.3;  $\text{CaCl}_2$ , 2-2.5. Picrotoxin (50  $\mu\text{M}$ ) was added to the bath



solution in ~60% of the experiments. Electrodes for whole-cell recording were filled with (in mM) caesium gluconate, 130; CsCl, 5; EGTA, 0.5;  $\text{MgCl}_2$ , 2; Na-ATP, 2; GTP, 0.2; NaCl, 5; HEPES 10; pH 7.25. When using the perforated patch technique, the pipette solution also contains 180  $\mu\text{g ml}^{-1}$  nystatin. E.p.s.cs were evoked by passing current through a bipolar tungsten electrode at 0.25 Hz while holding the neuron's membrane potential at  $-70$  mV. Tetanic stimulation consisted of three bursts, each at 100 Hz with a 200 ms duration and an interburst interval of 10 s, that were applied while the target neuron's membrane potential was maintained at  $-30$  mV. LTP was also induced by a pairing protocol with 15 stimuli applied at 0.2 Hz while maintaining the neuron's membrane potential at 0 mV.

FIG. 2 *a*, Average records of current as a function of time with current and time given by the calibration bars for all traces. The left traces were taken from a typical experiment during the time preceding the tetanus and the right-hand traces from a final 10-min period 30 min after the tetanus. The top traces are averages of synaptic currents (all events,  $N=110$  left,  $N=150$  right), the middle traces are averages of those records for which no release was detected (failure,  $N=72$  left,  $N=67$  right), and the bottom traces are averages of records for which transmitter release was detected (release,  $N=38$  left,  $N=83$  right). *b*, Synaptic strength as a function of time (min) during the experiment demonstrating LTP for a sample neuron. Average peak synaptic current, including failures, was found by averaging across non-overlapping groups of successive individual peak synaptic currents and is plotted as a ratio to the average of the five baseline points. The same time base is also used for *c* and *d*. Tetanic stimulation was presented, using the same stimulus strength, at the time indicated by the arrow (also in *c* and *d*). The horizontal line after the tetanic stimulation represents LTP of 2.18. *c*, Release failure rate, estimated over the same groups of successive individual traces as used for the average synaptic strength in *b*, as a function of time. The failure rate decreased from 0.6 to 0.2. *d*, Average amplitude of the peak synaptic current for those trials on which release was detected, as a function of time; averages over the same groups of successive traces as used in *b* and *c*. *e*, Histograms presenting the relative frequency of peak e.p.s.c. (when release occurred) before (bars,  $N=151$ ) and after stable LTP has been established (smooth line,  $N=339$ ). The two histograms are not significantly different (Kolmogorov-Smirnov test,  $P > 0.1$ ).

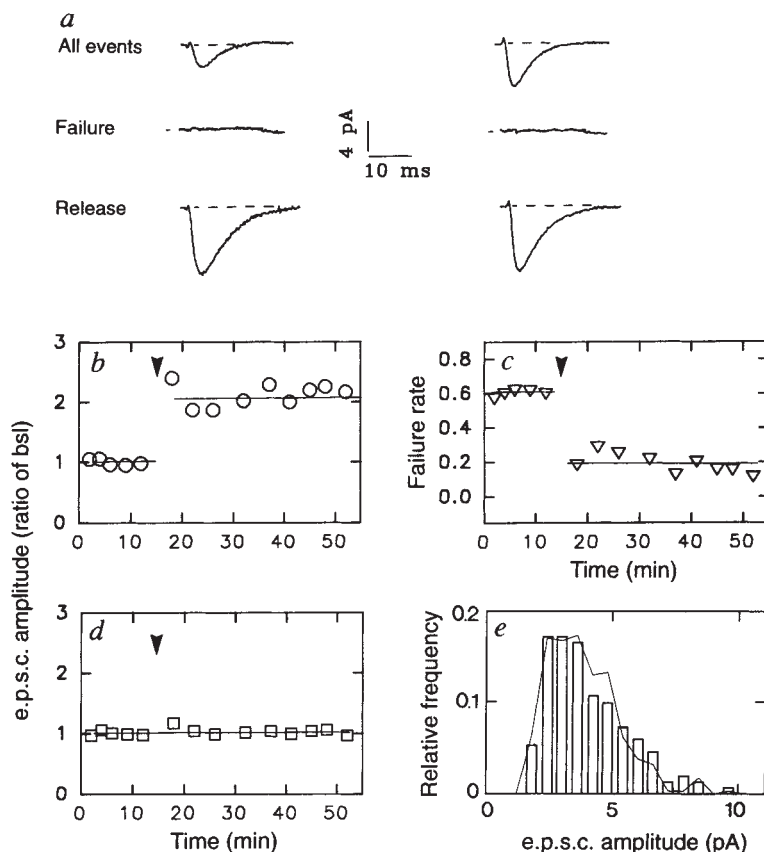




FIG. 3 *a*, Averages for a typical experiment as for Fig. 2*a*. All events,  $N=150$  left,  $N=180$  right; failure,  $N=68$  left,  $N=108$  right; release,  $N=82$  left,  $N=72$  right. *b*, Characteristics of responses for an LTD experiment plotted in the same way as the LTP data in Fig. 2. Low-frequency stimulation is indicated by bars in *b*, *c*, and *d*. LTD of 0.51 indicated by the line after the low-frequency stimulation. *c*, Failure rate increases from 0.6 to 0.77. *e*, Histogram of peak synaptic currents before low-frequency stimulation (open bars,  $N=80$ ) and after (solid line,  $N=100$ ). The amplitude histograms are not significantly different (Kolmogorov-Smirnov test,  $P>0.1$ ). METHODS. LTD was induced by applying stimuli at 1 Hz for 10–12 min while maintaining the membrane potential at  $-70$  mV. The magnitude of LTP or LTD was typically determined between 15 to 35 min after the conditioning event. Access resistance (15 to 30 M $\Omega$ ) and input resistance (100–300 M $\Omega$ ) were monitored throughout experiments and only recordings with stable access and input resistance were used for data analysis. A release success or failure was usually determined by eye, rather than by setting an arbitrary threshold, to ensure that any changes in the smallest response size were detected. A failure was determined by the following criteria: (1) the recorded trace is not discernibly different from background noise level; (2) the time course or shape of recorded trace is not consistent with the established time course or shape for the e.p.s.c.s from that neuron. In some cells, a small fraction of recorded traces (<5%) could not be fairly judged as either a failure or a success and were rejected. The distribution and average of e.p.s.c. amplitudes in our study are compatible with those of spontaneous miniature e.p.s.c. or single fibre e.p.s.c. in CA1 reported previously<sup>1,18</sup>. d-AP-5 (CRB), picrotoxin (Aldrich) and nystatin (Sigma).

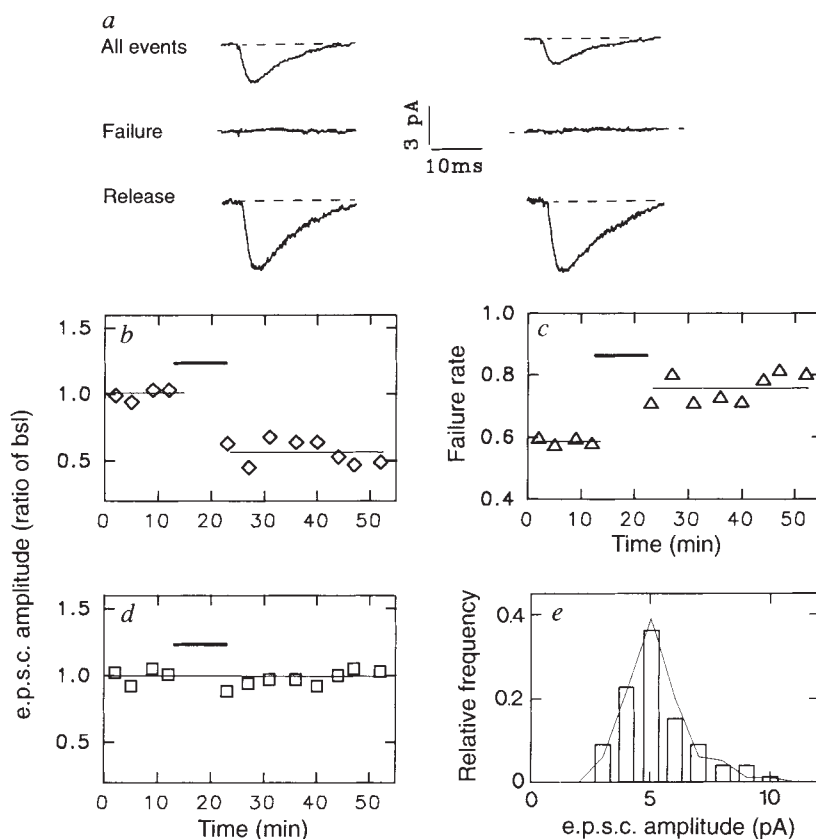
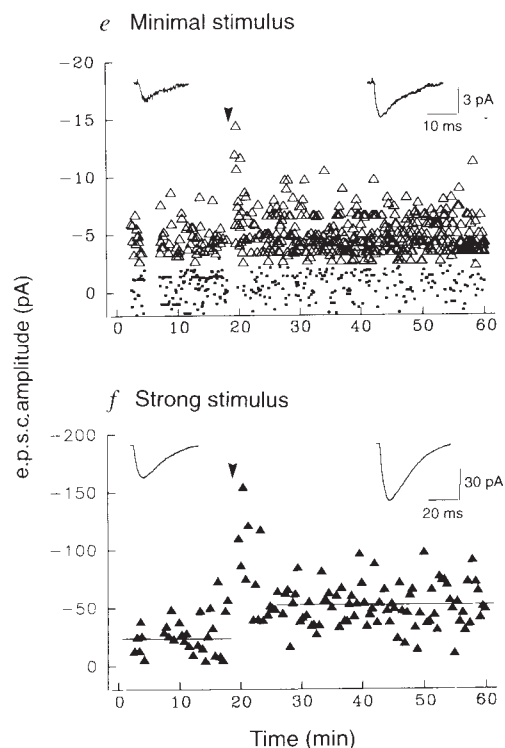
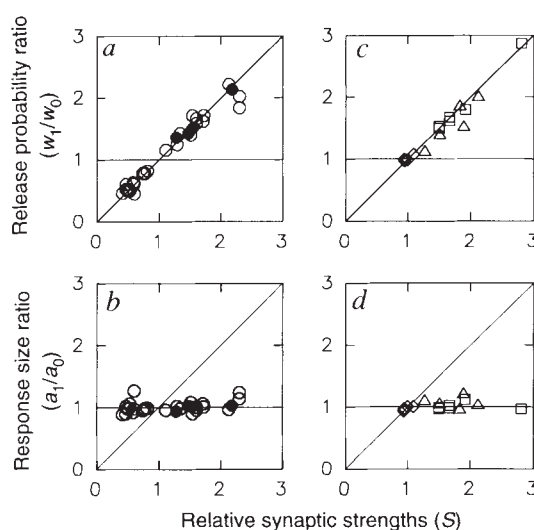


FIG. 4 *a*, Ratio of success probabilities (after/before conditioning stimulation consisting of tetanus, or low-frequency stimulation) as a function of synaptic strength ratio (after/before) calculated from averages such as illustrated in the top traces of Figs 2*a* and 3*a*. Unfilled circles are from 12 LTP (synaptic strength ratio  $>1$ ) and 14 LTD (synaptic strength ratio  $<1$ ) experiments. Filled circles are from 4 LTP experiments, of the sort illustrated in *e* and *f*, that interleaved strong and minimal stimuli at the same stimulation site in the slice. *b*, Ratio (after/before conditioning stimulation) of average peak e.p.s.c. for trials on which release occurred, obtained from averaged traces like those illustrated at the bottom of Figs 2*a* and 3*a*, as a function of the synaptic strength ratio (from average traces like those at the top of Figs 2*a* and 3*a*). Symbols have the same meanings as those in *a*. *c*, LTP experiments, as *a*. Diamonds are from 5 LTP experiments done in the presence of AP-5 ( $n=3$ ) or with the membrane potential maintained at  $-90$  mV during the tetanic stimulation ( $n=2$ ). Unfilled triangles are from 5 experiments in which 'pairing' rather than tetanic stimulation was the conditioning event. Two of these experiments used weak/strong conditioning stimuli (as in *e* and *f*). Unfilled squares are from six experiments that used perforated patch recording. *d*, Same type of plot as *b* with symbols having the same meaning as *c*. *e*, Peak synaptic current as a function of time during an LTP experiment with minimal stimulation. Standard tetanic stimulation with high intensity just before the 20 min time point. Insets are averages (over 10 min) of traces from before and 30 min after the tetanic stimulation, with current and time calibrations. Triangles from trials on which release was detected and dots for



no-release traces. *f*, As for *c*, but e.p.s.c.s evoked by strong stimulus applied to the same stimulating electrode. No release failures were seen with the strong stimulus. Note that the ordinate in *d* is 10 times greater than the ordinate in *c*. In these experiments, minimal and strong stimuli were applied alternatively to the stimulation site; the stimulus duration for minimal stimulation was reduced to 50  $\mu$ s and the stimulus duration for strong stimulation was 100  $\mu$ s with the same current as minimal stimulation.

of the postsynaptic cell during the conditioning event<sup>11,13,14</sup>. Data from five experiments that used hyperpolarization or AP-5 to block LTP, plotted as unfilled diamonds in Fig. 4c and d, all fall near  $S=1$ ,  $w_1/w_0=1$ , and  $a_1/a_0=1$ . Thus, LTP is produced under the conditions of our experiments by 'pairing' as well as tetanic stimulation, and LTP is blocked by hyperpolarization and AP-5.

Although we are aware of no firm evidence that genuine LTP requires the activation of a synaptic population, one can imagine some intersynaptic cooperative effect for which plasticity mechanisms would be different from those that are invoked when only one or two synapses are involved, as is the case for our preceding experiments. To test this, we have done the following experiment on six cells. At a single stimulation site, microscopic and macroscopic stimuli are interleaved with a pattern of five minimal stimuli (failure rate  $>0.5$ ) followed by one strong stimulus. After obtaining a baseline response size with this repeating 5/1 pattern, LTP was produced with the strong stimulus and then the 5 weak/1 strong stimulation pattern was resumed. Typical LTP was produced in the population of synapses activated by the strong stimulus, and LTP as we have measured it above was produced in the subset of synapses activated by the minimal stimulus (Fig. 4e,f).

Data from these experiments (filled circles for tetanic stimulation in Fig. 4a, b, and two unfilled triangles for 'pairing' in Fig. 4c and d) show that when LTP is elicited in a population of synapses, the modification of synaptic strength in the subset of synapses activated by the weak stimulus occurs by a change in synaptic reliability not potency.

All of the experiments presented above used standard whole-cell recording which can gradually change the concentration of small intracellular molecules; for example, LTP is known to 'wash out' after about 20 min. Could we have washed out some-

thing that is required for increasing potency? To test this, we did six LTP experiments using the perforated patch technique<sup>15</sup> and found that, as for the other experiments, LTP involves a change in reliability without modification of potency (squares in Fig. 4c and d).

Although many different mechanisms are used to modify synaptic strength, we conclude that, under the conditions of our experiments, LTP and LTD operate by modifying synaptic reliability, not potency. Whatever the precise mechanism by which it is achieved, the fact that synapses are so unreliable and that synaptic plasticity reflects modifications in reliability rather than potency has implications for information processing. The brain must use a design that is fault tolerant and produces reliable results when constructed with unreliable communications links that modify the strength of connections by changing reliability and not potency. □

Received 25 April; accepted 31 August 1994.

1. Raastad, M., Storm, J. F. & Andersen, P. *Eur. J. Neurosci.* **4**, 113–117 (1992).
2. Rosenmund, C., Clements, J. D. & Westbrook, G. L. *Science* **262**, 754–757 (1993).
3. Hessler, N. A., Shirker, A. M. & Malinow, R. *Nature* **366**, 569–572 (1993).
4. Allen, C. & Stevens, C. F. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
5. Bliss, T. V. P. & Collingridge, G. L. *Nature* **361**, 31–39 (1993).
6. Dudek, S. M. & Bear, M. F. *Proc. natn. Acad. Sci. U.S.A.* **89**, 4363–4367 (1992).
7. Mulkey, R. M. & Malenka, R. C. *Neuron* **9**, 967–975 (1992).
8. Edwards, F. A., Konnerth, A., Sakmann, B. & Takahashi, T. *Pflügers Arch. Ges. Physiol.* **414**, 600–612 (1989).
9. Blanton, M. G., Lo Turco, J. J. & Kriegstein, A. R. *J. Neurosci.* **30**, 203–210 (1989).
10. Malinow, R. & Tsien, R. W. *Nature* **346**, 177–180 (1990).
11. Wigström, H., Gustafsson, B., Huang, Y.-Y. & Abraham, W. C. *Acta Physiol. Scand.* **126**, 317–319 (1986).
12. Collingridge, G. L., Kehrl, S. J. & McLennan, H. J. *Physiol., Lond.* **334**, 33–46 (1983).
13. Malinow, R. & Miller, J. P. *Nature* **320**, 529–530 (1986).
14. Kelso, S. R., Ganong, A. H. & Brown, T. H. *Proc. natn. Acad. Sci. U.S.A.* **83**, 5326–5330 (1986).
15. Korn, S. J., Marty, A., Connor, J. A. & Horn, R. *Meth. Neurosci.* **4**, 364–372 (1991).
16. Stevens, C. F. & Wang, Y. *Nature* **364**, 147–149 (1993).
17. Silva, A. J., Stevens, C. F., Tonegawa, S. & Wang, Y. *Science* **257**, 201–206 (1992).
18. Manabe, T., Renner, P. & Nicoll, R. A. *Nature* **355**, 50–55 (1992).

## Cloning of an avermectin-sensitive glutamate-gated chloride channel from *Caenorhabditis elegans*

Doris F. Cully, Demetrios K. Vassilatis\*†, Ken K. Liu, Philip S. Pareiss, Lex H. T. Van der Ploeg\*, James M. Schaeffer & Joseph P. Arena

Department of Cellular Biochemistry and Physiology and  
\*Department of Genetics and Molecular Biology,  
Merck Research Laboratories, Rahway, New Jersey 07065-0900, USA

THE avermectins are a family of macrocyclic lactones used in the control of nematode and arthropod parasites<sup>1</sup>. Ivermectin (22,23-dihydroavermectin B<sub>1a</sub>) is widely used as an anthelmintic in veterinary medicine and is used to treat onchocerciasis or river blindness in humans<sup>1,2</sup>. Abamectin (avermectin B<sub>1a</sub>) is a miticide and insecticide used in crop protection<sup>1</sup>. Avermectins interact with vertebrate and invertebrate GABA receptors<sup>3–7</sup> and invertebrate glutamate-gated chloride channels<sup>8–11</sup>. The soil nematode *Caenorhabditis elegans* has served as a useful model to study the mechanism of action of avermectins<sup>11–15</sup>. A *C. elegans* messenger RNA expressed in *Xenopus* oocytes encodes an avermectin-sensitive glutamate-gated chloride channel<sup>11,14</sup>. To elucidate the structure and properties of this channel, we used *Xenopus* oocytes for expres-

sion cloning of two functional complementary DNAs encoding an avermectin-sensitive glutamate-gated chloride channel. We find that the electrophysiological and structural properties of these proteins indicate that they are new members of the ligand-gated ion channel superfamily.

*Xenopus* oocytes injected with *C. elegans* poly(A)<sup>+</sup> RNA exhibit a rapidly activating reversible glutamate- and irreversible ivermectin 4'-O-phosphate (IVMPO<sub>4</sub>)-sensitive current (Fig. 1a)<sup>11</sup>. A directional cDNA library was constructed from size-fractionated RNA and screened by expression in *Xenopus* oocytes. Glutamate- and IVMPO<sub>4</sub>-sensitive currents were observed after injection of *in vitro* RNA from a pool of 5,000 cDNAs (Fig. 1a). Subfractionation yielded two cDNA clones (pGluCl- $\alpha$  and pGluCl- $\beta$ ) that expressed glutamate- and IVMPO<sub>4</sub>-sensitive currents (Fig. 1a).

Oocytes simultaneously expressing GluCl  $\alpha$  and  $\beta$  proteins exhibited rapidly activating reversible glutamate- and irreversible IVMPO<sub>4</sub>-sensitive currents comparable with those found in oocytes injected with poly(A)<sup>+</sup> RNA (Fig. 1a). The time for maximal activation of IVMPO<sub>4</sub>-sensitive current was  $42 \pm 2$  s for GluCl  $\alpha$  and  $\beta$  compared to  $36 \pm 3$  s for poly(A)<sup>+</sup> RNA. The desensitization of the glutamate-sensitive current observed in oocytes injected with poly(A)<sup>+</sup> RNA was also observed in oocytes injected with GluCl  $\alpha$  and  $\beta$  at glutamate concentrations greater than 1 mM. The individual subunits, GluCl- $\alpha$  or GluCl- $\beta$ , expressed functional homomeric channels which were selectively responsive to IVMPO<sub>4</sub> or glutamate, respectively (Fig. 1a). The time course for IVMPO<sub>4</sub> activation of homomeric GluCl- $\alpha$  channels was  $18 \pm 1$  s, which is faster than that observed for GluCl  $\alpha$  and  $\beta$  and poly(A)<sup>+</sup> RNA ( $P < 0.001$ ).

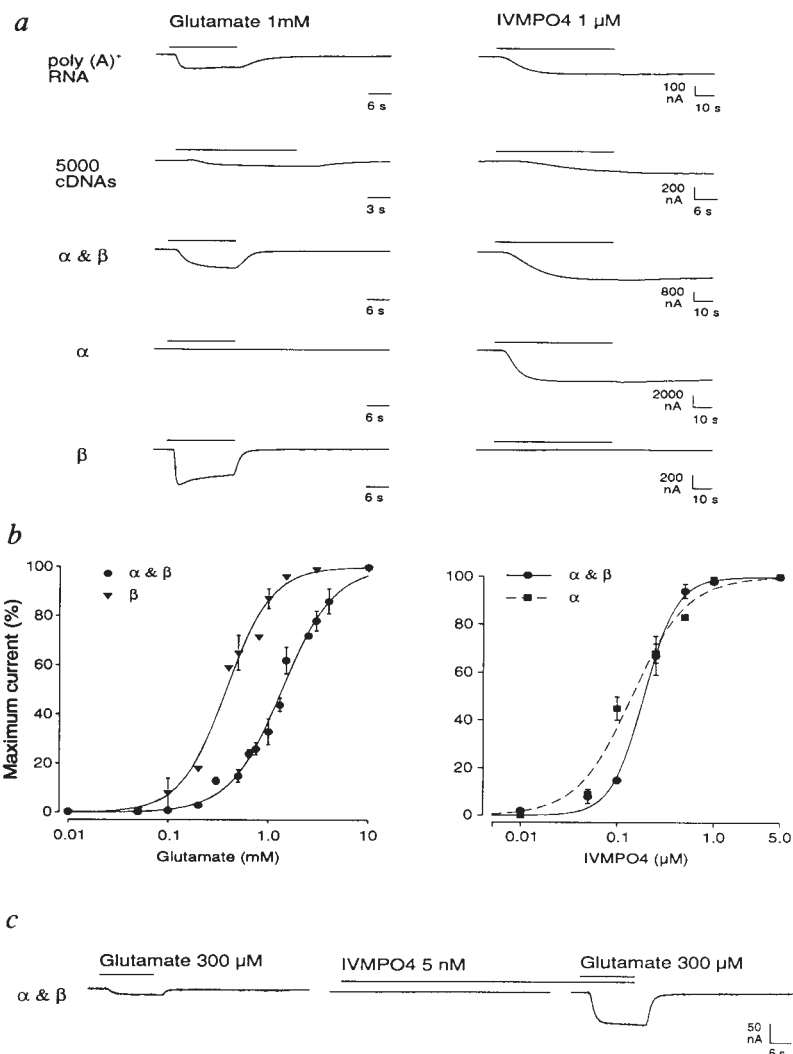
The dose-response curves indicate that coexpression of GluCl  $\alpha$  and  $\beta$  result in changes in the effector concentration for half-maximum response (EC<sub>50</sub>) for glutamate and in the Hill

† Visiting scientist from the Department of Genetics and Development, Columbia University, New York City, New York, USA.



FIG. 1 Electrophysiological characterization of the *C. elegans* glutamate- and IVMPO<sub>4</sub>-sensitive currents in *Xenopus* oocytes. Oocytes were injected with the following RNAs: *C. elegans* poly(A)<sup>+</sup> RNA (50 ng); or *in vitro* RNA from a pool of 5,000 cDNAs (50 ng), from pGluCl  $\alpha$  and  $\beta$  ( $\alpha$  &  $\beta$ ; 25 pg each), and from pGluCl- $\alpha$  or pGluCl- $\beta$  ( $\alpha$ ,  $\beta$ ; 250 pg). **a**, Expression of glutamate- and IVMPO<sub>4</sub>-sensitive currents. Current was recorded in the presence of glutamate and 1 min later in the presence of IVMPO<sub>4</sub>. Note the difference in timescale for the 5,000 cDNA current traces. **b**, Concentration-response curves for glutamate and IVMPO<sub>4</sub> in oocytes injected with *in vitro* RNA from pGluCl- $\alpha$  and pGluCl- $\beta$ , as indicated. Maximal currents were obtained with 10 mM glutamate or 3  $\mu$ M IVMPO<sub>4</sub>. The EC<sub>50</sub> and Hill coefficient were determined by fitting to the Michaelis-Menten equation (smooth curves)<sup>11,14</sup>. **c**, Modulation of glutamate-sensitive current by IVMPO<sub>4</sub>. Compounds were applied sequentially, at the indicated concentrations, with a 1-min lag time between current traces.

**METHODS.** Two functional cDNA clones for the avermectin and glutamate-sensitive chloride channel were isolated as follows: RNA was prepared from mixed stages of *C. elegans*<sup>14</sup> denatured in 80% formamide, 0.1% SDS and 1 mM EDTA for 10 min at 65 °C, and electrophoresed on a 2% SeaPrep (FMC) agarose gel (buffer system: 15 mM sodium phosphate, 1 mM EDTA pH 6.5). A 1.8-kilobase (kb) RNA fraction, giving electrophysiological responses to IVMPO<sub>4</sub> and glutamate, was used to synthesize a unidirectional cDNA library. First-strand cDNA was synthesized with the oligonucleotide primer (5'-(GA)<sub>10</sub>GCGGCCGC(T)<sub>18</sub>-3')<sup>26</sup> and second-strand cDNA synthesized as described<sup>27</sup>. DNA was treated with T4 DNA polymerase, NotI-digested and ligated with EcoRV-NotI-digested pBluescript SKII(+) (Stratagene). DNA was electroporated into TOP10 *E. coli* cells (Invitrogen) and clones selected by growth in liquid media with ampicillin. DNA extracted from amplified libraries representing  $2 \times 10^5$  recombinants was NotI-digested and size-separated by gel electrophoresis. The 4.0–4.7-kb linear DNA was recovered and recloned in pools of 5,000 recombinants. RNA was synthesized *in vitro*<sup>28</sup>, injected into oocytes, and glutamate- and IVMPO<sub>4</sub>-sensitive currents were observed. One positive pool of 5,000 cDNAs was grouped into smaller pools of 500 cDNAs, and two pools were identified which, when combined, expressed a glutamate- and IVMPO<sub>4</sub>-sensitive current. Currents were not observed in oocytes injected with RNA from the individual pools of 500 cDNAs. The cDNA pGluCl- $\alpha$  was isolated by subfractionating one of the pools of 500 cDNAs, transcribing RNA *in vitro*, and screening oocytes coinjected with *in vitro* RNA from the second pool of 500 cDNAs. The second pool of 500 cDNAs was similarly subfractionated and screened by coinjection with *in vitro* RNA from pGluCl- $\alpha$ , resulting in the isolation of the cDNA pGluCl- $\beta$ . When the complexity of the pools was reduced to 25 cDNAs, glutamate- and/or IVMPO<sub>4</sub>-sensitive current was observed without coinjection. *Xenopus* oocytes were prepared, injected and



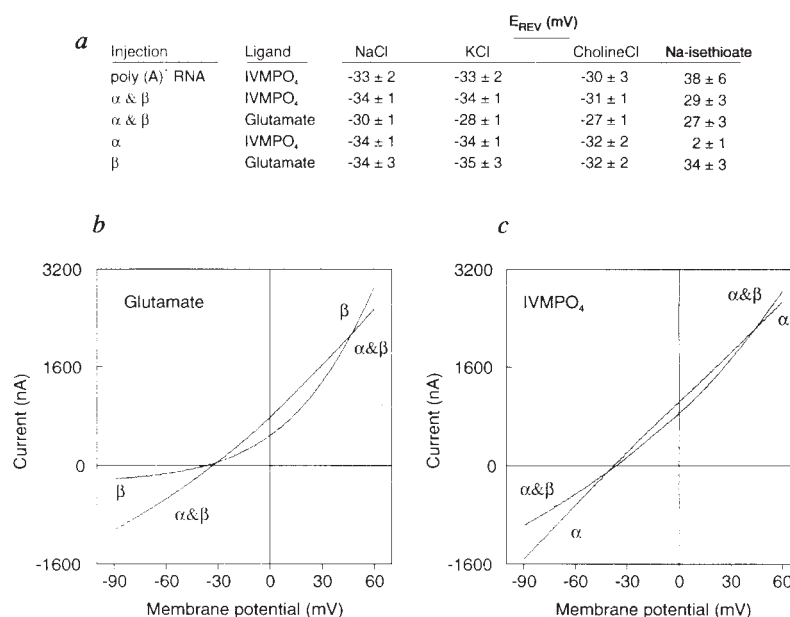
membrane currents recorded as previously described<sup>14</sup>. Recordings were made 2–10 days after injection. Unless otherwise indicated, recordings were made in standard frog saline consisting of (in mM): NaCl, 115; KCl, 2; MgCl<sub>2</sub>, 1; CaCl<sub>2</sub>, 1.8; HEPES, 10, adjusted to pH 7.5 with NaOH at 22 °C at –80 mV. The water-soluble avermectin derivative, IVMPO<sub>4</sub>, was used for these studies<sup>14</sup>. The bath was connected to ground with Ag/AgCl electrode directly, or through a 3 M KCl agar bridge when extracellular chloride was varied. The perfusion system had a dead time of 1 s; solid lines above the current traces indicate the time for which the perfusion system was switched to drug-containing solution. All data are means  $\pm$  s.e.m.

coefficient for IVMPO<sub>4</sub> (Fig. 1b). The coexpression resulted in a shift in the EC<sub>50</sub> for glutamate from  $380 \pm 20$  to  $1,360 \pm 50$   $\mu$ M. The Hill coefficients of  $1.9 \pm 0.2$  for GluCl- $\beta$  and  $1.7 \pm 0.1$  for GluCl  $\alpha$  and  $\beta$  suggested that more than one glutamate molecule was necessary to gate the channels. The EC<sub>50</sub> for IVMPO<sub>4</sub> activation of current was similar in oocytes injected with GluCl- $\alpha$  and with GluCl  $\alpha$  and  $\beta$ , with values of  $140 \pm 15$  and  $190 \pm 7$  nM, respectively (Fig. 1c). The Hill coefficient for IVMPO<sub>4</sub> was altered from  $1.5 \pm 0.2$  for GluCl- $\alpha$  to  $2.5 \pm 0.2$  for GluCl  $\alpha$  and  $\beta$ .

IVMPO<sub>4</sub> has a dual effect on oocytes injected with *C. elegans* poly(A)<sup>+</sup> RNA<sup>11</sup>. In addition to direct activation of current (Fig. 1a), the glutamate-sensitive current is potentiated by low concentrations of IVMPO<sub>4</sub><sup>11</sup>. Likewise, IVMPO<sub>4</sub> (5 nM) potentiated the glutamate-sensitive current  $490 \pm 45\%$  in oocytes injected with GluCl  $\alpha$  and  $\beta$ . The glutamate response from homomeric GluCl- $\beta$  channels was not potentiated with IVMPO<sub>4</sub> (data not shown).

The permeability properties and current-voltage relationship in oocytes expressing GluCl  $\alpha$  and  $\beta$  channels were also similar to those observed in oocytes injected with poly(A)<sup>+</sup> RNA (Fig. 2)<sup>11,14</sup>. The GluCl  $\alpha$  and  $\beta$  channels were selective for chloride, as shown by the shift in the reversal potential for glutamate- or IVMPO<sub>4</sub>-sensitive currents after replacement of external NaCl with sodium isethionate (Fig. 2a). The GluCl  $\alpha$  and  $\beta$  channels were not permeable to monovalent cations; replacement of external NaCl with KCl or choline-Cl did not shift  $E_{rev}$  (Fig. 2a). Permeability studies for homomeric GluCl- $\alpha$  or GluCl- $\beta$  channels also revealed selectivity for anions over cations (Fig. 2a). Replacement of NaCl with sodium isethionate shifted the  $E_{rev}$  to positive voltages, whereas replacement with KCl or choline-Cl had no effect (Fig. 2a). The  $E_{rev}$  of GluCl- $\alpha$  channels in sodium isethionate indicated that there was permeability to the large anion isethionate with a ratio to chloride of 0.2 (refs 16, 17).

FIG. 2 Permeability and voltage dependence. Oocytes were injected with poly(A)<sup>+</sup> RNA (50 ng), *in vitro* RNA from pGluCl  $\alpha$  and  $\beta$  ( $\alpha$  &  $\beta$ , 25 pg each) or individually with 250 pg of pGluCl- $\alpha$  ( $\alpha$ ) or pGluCl- $\beta$  ( $\beta$ ) as indicated. Currents were elicited with a 1 s voltage ramp from -90 to 60 mV in the absence and presence of glutamate (1 mM) or IVMP<sub>4</sub> (1  $\mu$ M) as indicated. Drug-sensitive currents were generated by subtracting drug-free from drug-containing data. a, Reversal potential ( $E_{REV}$ ) for IVMP<sub>4</sub>- and glutamate-sensitive currents obtained under the conditions indicated. Ion substitutions were accomplished by replacing NaCl in standard frog saline with the salts indicated. b, Glutamate-sensitive  $I/V$  relationships obtained in frog saline from oocytes injected with GluCl- $\beta$  ( $\beta$ ) or GluCl  $\alpha$  and  $\beta$  ( $\alpha$  &  $\beta$ ). c, IVMP<sub>4</sub>-sensitive  $I/V$  relationships obtained in frog saline from oocytes injected with GluCl  $\alpha$  ( $\alpha$ ) or GluCl  $\alpha$  and  $\beta$  ( $\alpha$  &  $\beta$ ).



The glutamate- and IVMP<sub>4</sub>-sensitive current-voltage relationships ( $I/V$ ) in oocytes injected with GluCl  $\alpha$  and  $\beta$  were outwardly rectifying, as previously observed for poly(A)<sup>+</sup> RNA (Fig. 2b, c)<sup>11,14</sup>. The  $I/V$  curves for the homomeric GluCl- $\alpha$  or GluCl- $\beta$  channels deviated strongly from the  $I/V$  curve for GluCl  $\alpha$  and  $\beta$ . Glutamate-sensitive current from GluCl- $\beta$  channels exhibited a steep outwardly rectifying voltage dependence (Fig. 2b), whereas the IVMP<sub>4</sub>-sensitive current from GluCl- $\alpha$  channels showed an essentially linear voltage dependence (Fig. 2c).

The pharmacological profile of oocytes injected with GluCl  $\alpha$  and  $\beta$  is distinct from other cloned ligand-gated chloride channels and glutamate-gated cation channels (Table 1). Ibotenate, a structural analogue of glutamate known to activate glutamate-gated chloride channels<sup>11,18,19</sup>, was the only ligand-gated ion channel agonist found to activate current (Table 1a). Glutamate- and IVMP<sub>4</sub>-sensitive currents were blocked by picrotoxin and flufenamic acid at concentrations similar to those reported for *C. elegans* poly(A)<sup>+</sup> RNA (Table 1b)<sup>11,14</sup>.

The nematocidal activity of avermectin analogues correlates with their ability to activate and potentiate glutamate-sensitive current in oocytes injected with poly(A)<sup>+</sup> RNA<sup>20</sup>. Four avermectin analogues, with different nematocidal activities in *C. elegans*, were tested on oocytes expressing GluCl  $\alpha$  and  $\beta$  (Table 1c). The order of nematocidal activity for these analogues is: ivermectin  $\approx$  milbemycin D > L648,548 > octahydroAVM<sup>13,20</sup>, which agrees with their ability to potentiate glutamate-sensitive current, and to activate IVMP<sub>4</sub>-sensitive current (Table 1c). These correlations suggest that the GluCl  $\alpha$  and  $\beta$  channel represents a target of avermectin in nematodes.

The nucleotide sequences of pGluCl- $\alpha$  and pGluCl- $\beta$  revealed single large open reading frames of 1,383 and 1,302 base pairs, respectively. The cDNAs have 5'- and 3'-untranslated regions of 50 and 90 nucleotides for pGluCl- $\alpha$ , and 13 and 144 nucleotides for pGluCl- $\beta$ . The nucleotide sequence in the regions where the open reading frames are located shared 50% identity. The first in-frame methionines were designated as initiation codons for open reading frames that predict a GluCl- $\alpha$  protein with an estimated  $M_r$  of 52,550 and a GluCl- $\beta$  protein with an estimated  $M_r$  of 49,900 (Fig. 3). Both putative proteins contained hydrophobic amino-terminal residues with sequences highly predictive of signal cleavage sites which would result in mature proteins initiating at amino-acid residue 21 in GluCl- $\alpha$  and 23

TABLE 1 Pharmacology

| (a) Agonists       |                               |                                      |                   |
|--------------------|-------------------------------|--------------------------------------|-------------------|
| Compound           | Concentration<br>( $\mu$ M)   | Per cent of glutamate<br>(10 mM)     |                   |
| IVMPO <sub>4</sub> | 1                             | 117 $\pm$ 13                         |                   |
| Ibotenate          | 500                           | 18 $\pm$ 2                           |                   |
| D-glutamate        | 10                            | 1 $\pm$ 2                            |                   |
| (b) Antagonists    |                               |                                      |                   |
| Compound           | Concentration<br>( $\mu$ M)   | Ligand                               | Inhibition<br>(%) |
| Picrotoxin         | 100                           | Glutamate 1 mM                       | 68 $\pm$ 4        |
| Picrotoxin         | 100                           | IVMPO <sub>4</sub> 1 $\mu$ M         | 61 $\pm$ 2        |
| Flufenamic<br>acid | 200                           | IVMPO <sub>4</sub> 1 $\mu$ M         | 60 $\pm$ 4        |
| Strychnine         | 100                           | Glutamate 1 mM                       | 0                 |
| Bicuculline        | 100                           | Glutamate 1 mM                       | 0                 |
| CNQX               | 10                            | Glutamate 1 mM                       | 0                 |
| (c) Avermectins    |                               |                                      |                   |
| Compound           | Potential of glutamate<br>(%) | IVMPO <sub>4</sub> activation<br>(%) |                   |
| Ivermectin         | 582 $\pm$ 113                 | 49 $\pm$ 10                          |                   |
| Milbemycin D       | 499 $\pm$ 155                 | 60 $\pm$ 4                           |                   |
| L648548            | 298 $\pm$ 38                  | 10 $\pm$ 3                           |                   |
| OctahydroAVM       | 70 $\pm$ 15                   | Inactive                             |                   |

Oocytes were injected simultaneously with GluCl  $\alpha$  and  $\beta$  (25 pg each). *n*, at least four for each group. a, The following compounds were inactive: NMDA, kainate, quisqualate, AMPA, acetylcholine and muscimol, tested at 1 mM; and L-aspartate, GABA, glycine, histamine,  $\beta$ -alanine and taurine, tested at 10 mM. The response to the indicated ligands was normalized to the response elicited with 10 mM glutamate. NMDA, *N*-methyl-D-aspartate; AMPA,  $\alpha$ -amino-3-hydroxyl-5-methyl-4-isoxazole propionic acid. b, Inhibition was considered to be zero if the response in the presence of antagonist was 0 ± 3% of control. The methochloride salt of bicuculline was used. CNQX, 6-cyano-7-nitroquinoline-2,3-dione. c, Per cent potentiation was calculated from the response to 300  $\mu$ M glutamate before and after a 2-min incubation with 5 nM of the analogue. The per cent activation was determined by activating current in the presence of analogue (500 nM) for 2 min, followed by maximal activation with 1  $\mu$ M IVMP<sub>4</sub>. Milbemycin D, 13-desoxy-22,23-dihydroavermectin B<sub>1b</sub> aglycone; L-648548, 22,23-dihydro-13-O-methoxyethoxymethylavermectin B<sub>1a</sub> aglycone; octahydroAVM, 3,4,8,9,10,11,22,23-octahydroavermectin B<sub>1</sub>.



The data show that the two *C. elegans* cDNA clones, pGluCl $\alpha$  and pGluCl $\beta$ , when expressed in *Xenopus* oocytes, form a chloride channel sensitive to glutamate, ibotenate, IVMPO $_4$  and a series of avermectin analogues. Each subunit can form distinct homomeric chloride channels. Several lines of evidence indicate that coexpression of GluCl  $\alpha$  and  $\beta$  leads to formation of heteromeric channels. Coexpression alters the time course of IVMPO $_4$  activation of current, the EC $_{50}$  for glutamate, the Hill coefficient

Sequence Analysis Software Package by Genetics Computer Group Inc. (GCG). Protein sequences were analysed with the GCG pileup, peptide structure and motif programs searching the Genbank, EMBL and Swis-Prot databases.

Glutamate-gated chloride channels have only been reported in invertebrates and are found on insect muscle and neuronal somata, crustacean muscle, and are expressed in oocytes injected with insect muscle poly(A)<sup>+</sup> RNA<sup>19,22–25</sup>. Like the *C. elegans* GluCl  $\alpha$  and  $\beta$  channels expressed in oocytes, arthropod glutamate-gated chloride channels are activated with ibotenate, blocked by high concentrations of picrotoxin, insensitive to GABA and are directly activated with avermectins<sup>8,18,19,23,24</sup>. Therefore, GluCl- $\alpha$  and GluCl- $\beta$  appear to represent a class of ligand-gated chloride channels closely related to arthropod glutamate-gated chloride channels. This class of channels represents a target of avermectins in *C. elegans*, and may mediate the paralyzing actions of avermectins in other organisms. □

1. Campbell, W. C. *Ivermectin and Abamectin* (Springer, New York, 1989).
2. Aziz, M. A., Diallo, S., Diop, I. M., Larièvre, M. & Porta, M. *Lancet* **2**, 171-173 (1982).
3. Fritz, L. C., Wang, C. C. & Gorio, A. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2062-2066 (1979).
4. Kass, I. S., Wang, C. C., Walrond, J. P. & Stretton, A. O. W. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6211-6215 (1980).
5. Pong, S.-S. & Wang, C. C. *J. Neurochem.* **38**, 375-379 (1982).
6. Mellin, T. N., Busch, R. D. & Wang, C. C. *Neuropharmacology* **22**, 89-96 (1983).
7. Sigel, E. & Baur, R. *Molec. Pharmac.* **32**, 749-752 (1987).
8. Scott, R. H. & Duce, I. R. *Pestic. Sci.* **16**, 599-604 (1985).
9. Duce, I. R. & Scott, R. H. *Brit. J. Pharmac.* **85**, 395-401 (1985).
10. Zufall, F., Franke, C. & Hatt, H. J. *exp. Biol.* **142**, 191-205 (1989).
11. Arena, J. P., Liu, K. K., Pareiss, P. S., Schaeffer, J. M. & Cully, D. F. *Molec. Brain Res.* **15**, 339-348 (1992).
12. Schaeffer, J. M. & Haines, H. W. *Biochem. Pharmac.* **38**, 2329-2338 (1989).
13. Cully, D. F. & Pareiss, P. S. *Molec. Pharmac.* **40**, 326-332 (1991).
14. Arena, J. P., Liu, K. K., Pareiss, P. S. & Cully, D. F. *Molec. Pharmac.* **40**, 368-374 (1991).

15. Rohrer, S. P., Meinke, P. T., Hayes, E. C., Mrozik, H. & Schaeffer, J. M. *Proc. natn. Acad. Sci. U.S.A.* **89**, 4168–4172 (1992).
16. Goldman, D. E. *J. gen. Physiol.* **27**, 37–60 (1943).
17. Hodgkin, A. L. & Katz, B. *J. Physiol., Lond.* **108**, 37–77 (1949).
18. Lea, T. J. & Usherwood, P. N. R. *Comp. gen. Pharmac.* **4**, 351–363 (1973).
19. Cull-Candy, S. G. *J. Physiol.* **255**, 449–464 (1976).
20. Arena, J. P. et al. *J. Parasitol.* (in the press).
21. Dingledine, R., Myers, S. J. & Nicholas, R. A. *FASEB J.* **4**, 2636–2645 (1990).
22. Lea, T. J. & Usherwood, P. N. R. *Comp. gen. Pharmac.* **4**, 333–350 (1973).
23. Lingle, C. & Marder, E. *Brain Res.* **212**, 481–488 (1981).
24. Wafford, K. A. & Sattelle, D. B. *J. exp. Biol.* **144**, 449–462 (1989).
25. Fraser, S. P. et al. *Molec. Brain Res.* **8**, 331–341 (1990).
26. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, New York, 1982).
27. Okayama, H. & Berg, P. *Molec. cell. Biol.* **2**, 161–170 (1982).
28. Julius, D., MacDermott, A., Axel, R. & Jessell, T. *Science* **241**, 558–564 (1988).

ACKNOWLEDGEMENTS. We thank P. Liberatore for synthesis of oligonucleotides.

## Long-term uncoupling of chloride secretion from intracellular calcium levels by $\text{Ins}(3,4,5,6)\text{P}_4$

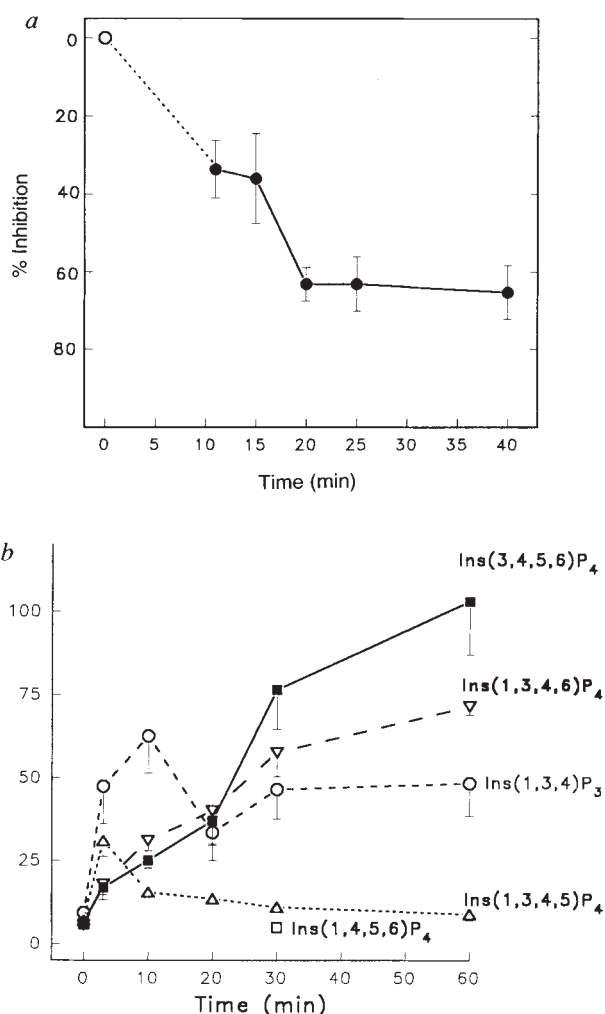
Mana Vajanaphanich, Carsten Schultz<sup>\*†</sup>, Marco T. Rudolf<sup>†</sup>, Matthew Wasserman, Péter Enyedi<sup>‡</sup>, Andrew Craxton<sup>‡</sup>, Stephen B. Shears<sup>‡</sup>, Roger Y. Tsien<sup>\*§</sup>, Kim E. Barrett & Alexis Traynor-Kaplan<sup>||</sup>

Departments of Medicine 8414 and \*Pharmacology, and §Howard Hughes Medical Institute, 0647, School of Medicine, University of California, San Diego, La Jolla, California 92093, USA  
<sup>†</sup>University of Bremen, Institute of Organic Chemistry, Bremen, Germany

<sup>‡</sup>Laboratory of Cellular and Molecular Pharmacology, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina 27709, USA

**OSMOREGULATION, inhibitory neurotransmission and pH balance depend on chloride ion ( $\text{Cl}^-$ ) flux. In intestinal epithelial cells, apical  $\text{Cl}^-$  channels control salt and fluid secretion and are, in turn, regulated by agonists acting through cyclic nucleotides and internal calcium ion concentration ( $[\text{Ca}^{2+}]_i$ )<sup>1–3</sup>. Recently, we found that muscarinic pretreatment prevents  $[\text{Ca}^{2+}]_i$  increases from eliciting  $\text{Cl}^-$  secretion in  $\text{T}_{84}$  colonic epithelial cells<sup>4</sup>. By studying concomitant inositol phosphate metabolism, we have now identified  $\text{D-myo-inositol 3,4,5,6-tetrakisphosphate}$  ( $\text{Ins}(3,4,5,6)\text{P}_4$ ), as the inositol phosphate most likely to mediate this uncoupling. A novel, membrane-permeant ester prepared by total synthesis delivers  $\text{Ins}(3,4,5,6)\text{P}_4$  intracellularly and confirms that this emerging messenger<sup>5</sup> does inhibit  $\text{Cl}^-$  flux resulting from thapsigargin- or histamine-induced  $[\text{Ca}^{2+}]_i$  elevations.**

In the human colonic epithelial cell line  $\text{T}_{84}$  (ref. 6), carbachol initially stimulates a transient elevation of  $[\text{Ca}^{2+}]_i$  and rise in  $\text{Cl}^-$  secretion which return to near-control levels within 10 min<sup>7,8</sup>. However, prolonged pretreatment with carbachol blocks subsequent thapsigargin-stimulated  $\text{Cl}^-$  secretion, without altering the ability of thapsigargin to increase  $[\text{Ca}^{2+}]_i$ . This does not appear to be mediated by protein kinase C (PKC)<sup>9</sup>. Similarly, we have now found that pretreatment with carbachol reduced  $10^{-4}\text{M}$  histamine-stimulated  $\text{Cl}^-$  secretion by an average of 77% (peak change in short circuit current ( $\Delta I_{sc}$ ) ( $\mu\text{A cm}^{-2}$ );  $n=3$ ,  $P<0.026$ ) without altering the  $[\text{Ca}^{2+}]_i$  response to histamine (ref. 8, and data not shown). Therefore, carbachol generates an intracellular signal capable of uncoupling the  $\text{Cl}^-$  secretory response from  $[\text{Ca}^{2+}]_i$ . We investigated the time course of this uncoupling. Carbachol-mediated inhibition of thapsigargin-stimulated  $\text{Cl}^-$



**FIG. 1** Comparison of the time course of the inhibition of  $\text{Ca}^{2+}$ -stimulated  $\text{Cl}^-$  secretion and carbachol-induced elevation of inositol phosphates in  $\text{T}_{84}$  cells. **a**, Carbachol-mediated inhibition of thapsigargin-stimulated  $\text{Cl}^-$  secretion relative to the 10 min, peak response to thapsigargin alone (open circle) is depicted. Carbachol ( $10^{-4}\text{M}$ ) was added to  $\text{T}_{84}$  cell monolayers mounted in Ussing chambers and, at subsequent intervals up to 30 min, thapsigargin ( $\text{Tg}$ ,  $10^{-6}\text{M}$ ) was added. After a further 10 min  $\text{Cl}^-$  secretion was assayed as the change in short circuit current ( $\Delta I_{sc}$   $\mu\text{A cm}^{-2}$ ) as described<sup>7</sup>. Data are means  $\pm$  s.e.m. for  $n=5$ . **b**, Carbachol ( $10^{-4}\text{M}$ ) was added to cells prelabelled with  $[\text{H}^3]\text{inositol}$  for 72 h and levels of inositol phosphates were measured at the indicated intervals. Data are c.p.m. expressed as a fraction of total cellular radioactivity and are means  $\pm$  s.e.m.,  $n=4$ . The techniques used to label the cells with  $[\text{H}^3]\text{inositol}$ , and the methods for quenching the cells, extracting the inositol phosphates, and resolving the isomers on Adsorbosphere SAX HPLC columns, are all as described previously<sup>4</sup>.  $[\text{H}^3]\text{Ins}(3,4,5,6)\text{P}_4$  and  $[\text{H}^3]\text{Ins}(1,4,5,6)\text{P}_4$  are enantiomers and elute from the column as a single peak. Their relative amounts were determined by incubating the peak with  $^{32}\text{P}\text{-Ins}(1,4,5,6)\text{P}_4$  and partially purified  $\text{Ins}(1,4,5,6)\text{P}_4$  3-kinase<sup>13</sup>. A comparison of zero-time peaks ( $56\% \pm 8\%$   $[\text{H}^3]\text{Ins}(1,4,5,6)\text{P}_4$ ,  $44 \pm 8\%$   $[\text{H}^3]\text{Ins}(3,4,5,6)\text{P}_4$ ,  $n=5$ ) with peaks that increased sevenfold after 30 min of stimulation with carbachol ( $5\text{--}7\%$   $[\text{H}^3]\text{Ins}(1,4,5,6)\text{P}_4$ ,  $93\text{--}95\%$   $[\text{H}^3]\text{Ins}(3,4,5,6)\text{P}_4$ ,  $n=2$ ) indicated that carbachol specifically increased levels of  $[\text{H}^3]\text{Ins}(3,4,5,6)\text{P}_4$  roughly 15-fold whereas levels of  $[\text{H}^3]\text{Ins}(1,4,5,6)\text{P}_4$  were unchanged. Similar changes were observed when  $\text{Ins}(3,4,5,6)\text{P}_4$  mass was measured directly (Fig. 4). Thus, levels of  $[\text{H}^3]\text{Ins}(3,4,5,6)\text{P}_4$  throughout the time course were calculated by subtracting the amount of  $[\text{H}^3]\text{Ins}(1,4,5,6)\text{P}_4$  at zero time from the other time points.

secretion was evident within 11 min, maximal by 20 min, and was maintained for more than 90 min (Figs 1a and 2a). The persistence of this phenomenon suggests the long-term effect of muscarinic agonists on  $\text{Cl}^-$  secretion will be inhibitory.

<sup>||</sup> To whom correspondence should be addressed.



Carbachol stimulates the formation of many inositol polyphosphates, any of which might be the uncoupling signal. Levels of a plausible candidate should rise and fall with a time course matching that of the inhibition of  $\text{Ca}^{2+}$ -dependent  $\text{Cl}^-$  secretion. The transient increases in  $[^3\text{H}]\text{Ins}(1,4,5)\text{P}_3$  and  $[^3\text{H}]\text{Ins}(1,3,4,5)\text{P}_4$  returned to baseline within 3 min<sup>4</sup> and 20 min (Fig. 1b), respectively, even while inhibition of  $\text{Cl}^-$  secretion was maximal (Fig. 1a). Levels of  $[^3\text{H}]\text{Ins}(1,4,5,6)\text{P}_4$  (Fig. 1b),  $[^3\text{H}]\text{Ins}(1,3,4,5,6)\text{P}_5$  (ref. 4) and  $[^3\text{H}]\text{InsP}_6$  (ref. 4) were not elevated after carbachol stimulation. Reversal of carbachol-mediated inhibition of thapsigargin-stimulated  $\text{Cl}^-$  secretion was delayed when prolonged carbachol stimulation preceded the addition of atropine (Fig. 2a). Thus, because levels of  $[^3\text{H}]\text{Ins}(1,3,4,6)\text{P}_4$  and  $[^3\text{H}]\text{Ins}(3,4,5,6)\text{P}_4$  and  $[^3\text{H}]\text{Ins}(1,3,4)\text{P}_3$  (which were all elevated at later time points corresponding to

inhibition of  $\text{Cl}^-$  secretion, Fig. 1b) returned to baseline at different rates following addition of atropine (Fig. 2b), it was possible to narrow the correlation further.  $\text{Ins}(3,4,5,6)\text{P}_4$  was the only inositol phosphate to remain elevated above basal levels while  $\text{Cl}^-$  secretion remained inhibited (Fig. 2). However, the

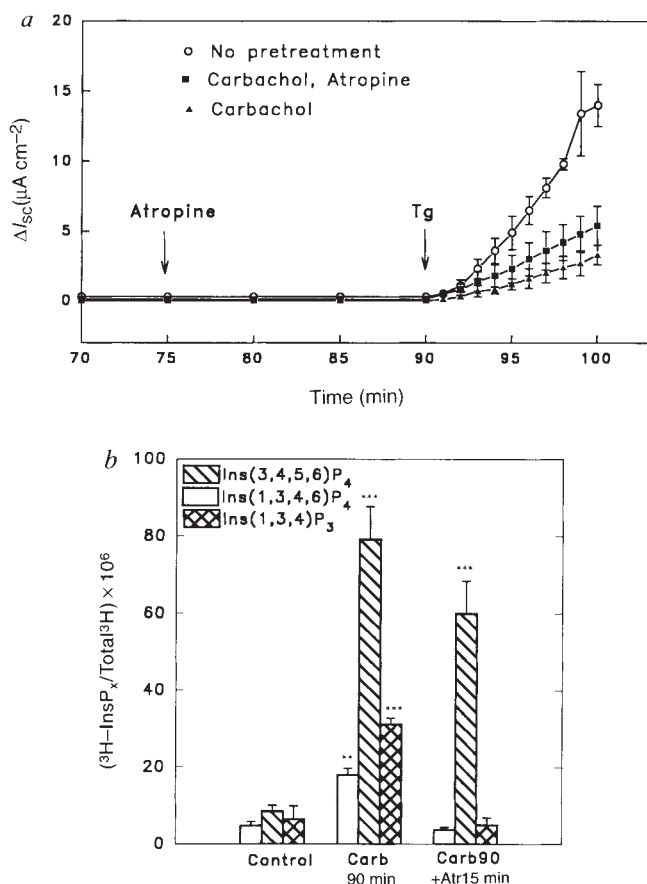


FIG. 2 Effect of atropine on carbachol-mediated increases in inositol phosphates and inhibition of  $\text{Ca}^{2+}$ -dependent  $\text{Cl}^-$  secretion. a, Time course of thapsigargin (Tg)-stimulated  $\text{Cl}^-$  secretion with no pretreatment or after pretreatment with  $10^{-6}$  M carbachol (added at time 0) with or without the addition of atropine ( $10^{-6}$  M) 15 min before Tg addition.  $\text{Cl}^-$  secretion across  $\text{T}_{84}$  monolayers mounted in Ussing chambers was monitored as short circuit current ( $I_{\text{sc}}$ ) as in Fig. 1. Data are means  $\pm$  s.e.m. of  $I_{\text{sc}}$  from individual monolayers,  $n=5$ . Monolayers were stimulated with carbachol for 90 min or carbachol for 75 min followed by atropine + carbachol for the remaining 15 min before the addition of Tg. Tg alone is shown as 'control'. b, Cells were labelled with  $^3\text{H}$ -inositol as described<sup>4</sup> and stimulated with carbachol ( $10^{-6}$  M) for 90 min. Atropine ( $10^{-6}$  M), where present, was added 75 min after carbachol. After an additional 15 min, cells were quenched and inositol phosphates were extracted and analysed by HPLC<sup>4</sup>. Data are c.p.m. expressed as fraction of total cellular radioactivity and are means  $\pm$  s.e.m.,  $n=6$ . The graph depicts levels of  $\text{Ins}(3,4,5,6)\text{P}_4$ ,  $\text{Ins}(1,3,4,6)\text{P}_4$  and  $\text{Ins}(1,3,4)\text{P}_3$  under the conditions indicated. All other inositol phosphates were at basal levels after 90 min carbachol + 15 min atropine. (Values that differ significantly from control values are denoted by asterisks: \*\* $P < 0.01$ ; \*\*\* $P < 0.001$  by Student's 2-tailed  $t$ -test).

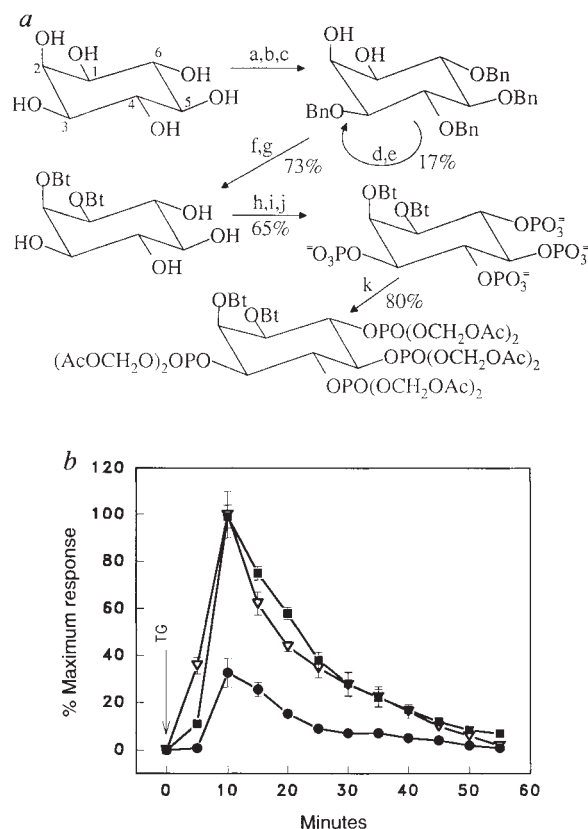


FIG. 3 Synthesis of  $\text{Bt}_2\text{Ins}(3,4,5,6)\text{P}_4/\text{AM}$  and  $\text{Bt}_2\text{Ins}(1,4,5,6)\text{P}_4/\text{AM}$  and effects on thapsigargin-stimulated  $\text{Cl}^-$  secretion. a, Synthesis of  $\text{Bt}_2\text{Ins}(1,4,5,6)\text{P}_4/\text{AM}$  and its racemic mixture with  $\text{Bt}_2\text{Ins}(3,4,5,6)\text{P}_4/\text{AM}$  starting from *myo*-inositol. Bt, *n*-butyryl; Bn, benzyl; Ac, acetyl. Reagents: a, cyclohexanone,  $\text{H}^+$ ; b, KOH, BnCl; c,  $\text{H}_3\text{O}^+$ ; d, camphanoyl chloride, pyridine, then crystallization and silica gel chromatography to separate the diastereomeric camphanates; e, KOH/MeOH; f,  $\text{Bt}_2\text{O}$ , pyridine; g,  $\text{H}_2/\text{Pd-C}$ ; h,  $\text{Et}_3\text{NP}(\text{OBn})_2$ , tetrazole; i,  $\text{CH}_3\text{CO}_3\text{H}$ ; j,  $\text{H}_2/\text{Pd-C}$ ; k, *i*-Pr<sub>2</sub>NEt,  $\text{AcOCH}_2\text{Br}$ . Steps a-c were as described in ref. 15. Yields for each group of subsequent steps are stated below the reaction arrows. The structures are drawn to show the stereochemistry of the active component,  $\text{Bt}_2\text{Ins}(3,4,5,6)\text{P}_4/\text{AM}$ , but the syntheses actually produced either its racemic mixture with the inactive enantiomer,  $\text{Bt}_2\text{Ins}(1,4,5,6)\text{P}_4/\text{AM}$  (when steps d and e were skipped) or the latter alone (in  $>99\%$  enantiomeric excess verified by chiral HPLC). b,  $\text{T}_{84}$  cell monolayers were treated with  $200 \mu\text{M}$   $\text{Bt}_2\text{Ins}(3,4,5,6)\text{P}_4/\text{AM}$  +  $200 \mu\text{M}$   $\text{Bt}_2\text{Ins}(1,4,5,6)\text{P}_4/\text{AM}$  (closed circles) or  $400 \mu\text{M}$   $\text{Bt}_2\text{Ins}(1,4,5,6)\text{P}_4/\text{AM}$  (closed squares) or buffer (open triangles) for 30 min in Ringers' solution at room temperature before  $10^{-6}$  M Tg-induced  $\text{Cl}^-$  secretion was measured as  $\Delta I_{\text{sc}}$ . Other control experiments showed that equivalent concentrations of  $\text{P}_4/\text{AM}$ ,  $\text{K}^+$  butyrate, or Pluronic/DMSO also were without effect. AM esters were dissolved in Pluronic F127/DMSO, final concentration 0.02%, 0.2%, respectively. Data measured in  $\mu\text{A cm}^{-2}$  were normalized to the maximum response obtained in the presence of Tg alone and are means  $\pm$  s.e.m. ( $n=2-4$ ) from a representative experiment. The average maximum response to Tg in monolayers pretreated with  $200 \mu\text{M}$   $\text{Bt}_2\text{Ins}(3,4,5,6)\text{P}_4/\text{AM}$  +  $200 \mu\text{M}$   $\text{Bt}_2\text{Ins}(1,4,5,6)\text{P}_4/\text{AM}$  was  $35.2 \pm 4.5\%$  of the control, (mean  $\pm$  s.e.m., 4 experiments in duplicate;  $P < 0.0001$ ). Neither  $\text{Bt}_2\text{Ins}(3,4,5,6)\text{P}_4/\text{AM}$  +  $\text{Bt}_2\text{Ins}(1,4,5,6)\text{P}_4/\text{AM}$  nor pure  $\text{Bt}_2\text{Ins}(1,4,5,6)\text{P}_4/\text{AM}$  altered resting levels of  $[\text{Ca}^{2+}]_i$ , or the Tg-mediated increase in  $[\text{Ca}^{2+}]_i$ , as assessed in Fura-2-loaded  $\text{T}_{84}$  cells grown on cover slips, nor the levels of endogenous inositol phosphates (data not shown).

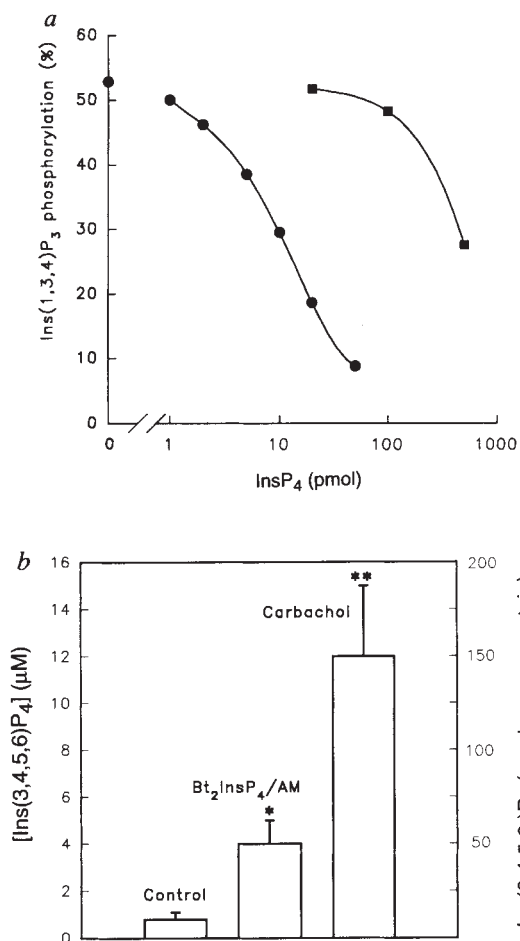


FIG. 4 Determination of levels of Ins(3,4,5,6)P<sub>4</sub> in T<sub>84</sub> cells treated with vehicle, carbachol or Bt<sub>2</sub>Ins(3,4,5,6)P<sub>4</sub>/AM. **a**, Standard curves are shown for the inhibition of partially purified Ins(1,3,4)P<sub>3</sub> 5/6-kinase<sup>16</sup> by either Ins(3,4,5,6)P<sub>4</sub> or Ins(1,4,5,6)P<sub>4</sub> (kindly supplied by S. Ozaki, Y. Watanabe and M. Hirata, Japan); data are means of duplicates (range <5% of mean) in the same experiment as data shown in **b**. These values were confirmatory of 2 previous experiments. ●, Ins(3,4,5,6)P<sub>4</sub>; ■, Ins(1,4,5,6)P<sub>4</sub>. Enzyme and inhibitor were incubated with 3 nCi [<sup>3</sup>H]Ins(1,3,4)P<sub>3</sub> (NEN-DuPont) for 30 min at 37 °C in 250 μl of medium containing 10 mM HEPES (pH 6.6), 6 mM MgSO<sub>4</sub>, 5 mM ATP, 10 mM creatine phosphate, 1.25 units creatine phosphokinase, 0.08 mg bovine serum albumin. Assays were quenched with perchloric acid, and neutralized with freon/octylamine<sup>17</sup> and analysed by HPLC<sup>4</sup>. Additional, duplicate 5/6-kinase assays contained extracts prepared from T<sub>84</sub> cells that had been incubated for 30 min with either 0.1 mM carbachol, 200 μM Bt<sub>2</sub>Ins(3,4,5,6)P<sub>4</sub>/AM + 200 μM Bt<sub>2</sub>Ins(1,4,5,6)P<sub>4</sub>/AM, or vehicle. These cells were quenched and extracted with phenol/chloroform<sup>18</sup>, spiked with 0.4 nCi [<sup>14</sup>C]Ins(1,3,4,5)P<sub>2</sub> (to monitor the HPLC resolution) + 3 nCi [<sup>3</sup>H]Ins(1,4,5,6)P<sub>4</sub> (obtained by 3-phosphatase mediated hydrolysis of [<sup>3</sup>H]Ins(1,3,4,5,6)P<sub>5</sub>, 1 Ci mmol<sup>-1</sup> (ref. 19)). 1-min fractions of HPLC eluates were collected, the Ins(3,4,5,6)P<sub>4</sub>/Ins(1,4,5,6)P<sub>4</sub> peak was located and desalted<sup>17</sup> and aliquots were incubated with 5/6-kinase. **b**, The mass of samples were measured as described in **a**. The proportion of Ins(1,3,4)P<sub>3</sub> phosphorylation yielded estimates of the mass levels of Ins(3,4,5,6)P<sub>4</sub>. The conversion to concentration was based on a T<sub>84</sub> cell volume of 10.7 μl per mg protein<sup>20</sup> (control, *n* = 4; 200 μM Bt<sub>2</sub>Ins(3,4,5,6)P<sub>4</sub>/AM + 200 μM Bt<sub>2</sub>Ins(1,4,5,6)P<sub>4</sub>/AM, *n* = 3; 10<sup>-4</sup> M carbachol, *n* = 3; Data are means ± s.e.m.; \**P* < 0.01).

inhibition of thapsigargin-induced Cl<sup>-</sup> secretion, and the elevation in [<sup>3</sup>H]Ins(3,4,5,6)P<sub>4</sub> levels by carbachol, were both reversed if a relatively short treatment with agonist (10 min) was accompanied by a prolonged treatment with atropine (50 min); this was the earliest point at which Ca<sup>2+</sup>-dependent Cl<sup>-</sup> secretion was fully restored (106 ± 6% of control values, *n* = 3), by which time [<sup>3</sup>H]Ins(3,4,5,6)P<sub>4</sub> levels had returned to 203 ± 49% of controls. For Ins(3,4,5,6)P<sub>4</sub> to mediate inhibition directly, we would expect its levels to be elevated at least threefold (Fig. 1). An eightfold increase produced full inhibition.

To confirm that Ins(3,4,5,6)P<sub>4</sub> regulates Cl<sup>-</sup> transport we directly introduced it into intact cells using cell permeant analogues. Although we were only able to prepare dibutyl-D-myo-inositol(3,4,5,6)P<sub>4</sub>/acetoxymethylester (Bt<sub>2</sub>Ins(3,4,5,6)P<sub>4</sub>/AM) as a racemic mixture that also contained Bt<sub>2</sub>Ins(1,4,5,6)P<sub>4</sub>/AM, we were able to synthesize enantiomerically pure Bt<sub>2</sub>Ins(1,4,5,6)P<sub>4</sub>/AM as a control. (See synthesis scheme, Fig. 3a). These compounds are cell-permeant because the charged phosphate groups are masked by acetoxymethyl ester (AM) groups<sup>11</sup>, whereas the hydroxyl groups are concealed by butyrates. Once inside the cell, the butyryl and acetoxymethyl ester groups are cleaved by intracellular esterases. Thapsigargin-mediated Cl<sup>-</sup> secretion was unaffected by 30 min pretreatment of T<sub>84</sub> cells with 400 μM Bt<sub>2</sub>Ins(1,4,5,6)P<sub>4</sub>/AM, or vehicle, but was substantially inhibited by 200 μM Bt<sub>2</sub>Ins(3,4,5,6)P<sub>4</sub>/AM + 200 μM Bt<sub>2</sub>Ins(1,4,5,6)P<sub>4</sub>/AM (Fig. 3). Thus, the effects are stereospecific for the 3,4,5,6-isomer and are not due to side products of ester hydrolysis. In the continued presence of the Bt<sub>2</sub>Ins(3,4,5,6)P<sub>4</sub>/AM, inhibition of Cl<sup>-</sup> transport lasted at least 45 min. The racemic mixture of 200 μM Bt<sub>2</sub>Ins(3,4,5,6)P<sub>4</sub>/AM + 200 μM Bt<sub>2</sub>Ins(1,4,5,6)P<sub>4</sub>/AM also inhibited 10<sup>-4</sup> M histamine-induced Ca<sup>2+</sup>-dependent Cl<sup>-</sup> secretion by 37.3 ± 4.8%

(mean ± s.e.m., *n* = 3; *P* < 0.01). The critical question of how much intracellular Ins(3,4,5,6)P<sub>4</sub> resulted from extracellular application of the permeant ester could not be answered by measurement of radioactivity, because labelling of the ester prepared as in Fig. 3a would require impractical amounts of radioactive inositol. Instead, a novel mass assay for Ins(3,4,5,6)P<sub>4</sub> was devised on the basis of its ability to inhibit Ins(1,3,4)P<sub>3</sub>-5/6-kinase *in vitro* (Fig. 4a). The results (Fig. 4b) depict a 15-fold stimulation by carbachol which agrees with values obtained from radioactive analysis (Fig. 1b). The basal level we report here is consistent with previous estimates of [Ins(3,4,5,6)P<sub>4</sub> + Ins(1,4,5,6)P<sub>4</sub>] in HL-60 cells<sup>12</sup>. The modest level of Ins(3,4,5,6)P<sub>4</sub> generated from the permeant ester (fivefold over baseline) may reflect some extracellular hydrolysis as well as competition between the 10 steps needed for complete de-esterification and the subsequent metabolism of the messenger.

Ins(3,4,5,6)P<sub>4</sub> probably acts without metabolic conversion, because its only significant transformation in T<sub>84</sub> cells (data not shown) and other tissues<sup>13</sup> is phosphorylation to Ins(1,3,4,5,6)P<sub>5</sub>. The latter isomer cannot be responsible because it already exists in unstimulated cells in very high quantities and these levels are relatively static<sup>4</sup>; furthermore, it can also be made from Ins(1,4,5,6)P<sub>4</sub><sup>13</sup>, yet Bt<sub>2</sub>Ins(1,4,5,6)P<sub>4</sub>/AM was biologically inert.

The lack of an exact correlation between the extent and duration of changes in [Ca<sup>2+</sup>]<sub>i</sub> and stimulation of Cl<sup>-</sup> secretion in T<sub>84</sub> cells has been puzzling<sup>8,10</sup>. Here we present a mechanism for this phenomenon by identifying conditions resulting in almost complete dissociation of Cl<sup>-</sup> secretion from [Ca<sup>2+</sup>]<sub>i</sub> levels, and by showing that this corresponds to an elevation of Ins(3,4,5,6)P<sub>4</sub> levels. In addition, we have demonstrated a sustained influence of agonists on Cl<sup>-</sup> transport long after the origi-



nal stimulus has ceased. This sustained, negative effect contrasts with the more transient and positive signalling responses usually attributed to alterations in inositol phosphate turnover. Furthermore, this effect of  $\text{Ins}(3,4,5,6)\text{P}_4$  exemplifies a new mechanism for desensitization of phospholipase C-linked responses<sup>14</sup>. Finally, the use of cell-permeant analogues demonstrate a novel means of studying inositol phosphate function in intact cells, which is technically less demanding than microinjecting single cells and is applicable to assays that require the use of cell populations, such as transepithelial secretion. □

Received 29 June; accepted 18 August 1994.

1. Frizzell, R. A., Welsh, M. J. & Smith, P. L. in *Ion Transport by Epithelia* (ed. Schultz, S. G.) Vol. 36, 137–145 (Soc. gen. Physiol. Series, Raven, New York, 1986).
2. Quinton, P. *FASEB J.* **4**, 2709–2717 (1990).
3. Dharmasathaphorn, K. & Traynor-Kaplan, A. E. in *Textbook of Internal Medicine* 2nd edn (ed. Kelley, W. N.) 398–404 (Lippincott, Philadelphia, 1991).
4. Kachintorn, U., Vajanaphanich, M., Barrett, K. E., Dharmasathaphorn, K. & Traynor-Kaplan, A. E. *Am. J. Physiol.* **264**, C671–C676 (1993).
5. Menniti, F. S., Oliver, K. G., Putney, J. W. Jr & Shears, S. B. *Trends biochem. Sci.* **18**, 53–56 (1993).

6. Dharmasathaphorn, K., McRoberts, J. A., Mandel, K. G., Tisdale, L. D. & Masui, H. *Am. J. Physiol.* **9**, G204–G208 (1984).
7. Dharmasathaphorn, K. & Pandol, S. J. *J. clin. Invest.* **77**, 348–354 (1986).
8. Dharmasathaphorn, K., Cohn, J. & Beuerlein, G. *Am. J. Physiol.* **256**, C1224–C1230 (1989).
9. Traynor-Kaplan, A. E., Buranawuti, T., Vajanaphanich, M. & Barrett, K. E. *Am. J. Physiol.* (in the press).
10. Kachintorn, U., Vajanaphanich, M., Traynor-Kaplan, A. E., Dharmasathaphorn, K. & Barrett, K. E. *Br. J. Pharmacol.* **109**, 510–517 (1993).
11. Schultz, C. et al. *J. Biol. Chem.* **268**, 6316–6322 (1993).
12. Pittet, D., Schlegel, W., Lew, D. P., Monod, A. & Mayr, G. W. *J. Biol. Chem.* **264**, 18489–18493 (1989).
13. Craxton, A., Erneux, C. & Shears, S. J. *Biol. Chem.* **269**, 4337–4342 (1994).
14. Wojcikiewicz, R. J. H., Tobin, A. B. & Nahorski, S. R. *Trends pharmac. Sci.* **14**, 279–285 (1993).
15. Angyal, S. J. & Tate, M. E. *J. chem. Soc.* **1965**, 6949–6955 (1965).
16. Abdullah, M. et al. *J. Biol. Chem.* **267**, 22340–22345 (1992).
17. Menniti, F. S., Miller, R. N., Putney, J. W. & Shears, S. B. *J. Biol. Chem.* **268**, 3850–3856 (1993).
18. Wong, N. S., Barker, C. J., Shears, S. B., Kirk, C. J. & Michell, R. H. *Biochem. J.* **252**, 1–5 (1988).
19. Nogimori, K. et al. *J. Biol. Chem.* **266**, 16499–16506 (1991).
20. Mandel, K. G., Dharmasathaphorn, K. & McRoberts, J. A. *J. Biol. Chem.* **261**, 704–712 (1986).

ACKNOWLEDGEMENTS. This work was supported in part by grants from the NIH and the AGA/Miles and Shirley Fiterman Foundation.

## Promoter-specific imprinting of the human insulin-like growth factor-II gene

Thanh H. Vu & Andrew R. Hoffman

Medical Service, VA Medical Center, and Department of Medicine, Stanford University, 3801 Miranda Avenue, Palo Alto, California 94304, USA

GENOMIC imprinting is a mechanism whereby only one of the two parental alleles is expressed. Loss or relaxation of genomic imprinting has been proposed as an epigenetic mechanism for oncogenesis in a variety of human tumours<sup>1–6</sup>. Although the mechanism of imprinting is unknown, differential CpG methylation of the parental alleles has been implicated<sup>7–12</sup>. The human insulin-like growth factor-II (*IGF2*) gene, which is transcribed from four promoters, P1–P4 (ref. 13), is imprinted in fetal liver<sup>14,15</sup> but biallelic expression occurs in adult liver<sup>16</sup>. Like most tissues, fetal liver uses primarily promoters P3 and P4 (ref. 17). Adult liver, however, transcribes *IGF2* from promoter P1, and it has been suggested that the recruitment of P1 may be responsible for the absence of imprinting in human liver, and in choroid plexus and leptomeninges<sup>18</sup>. We report here that in liver and chondrocytes, *IGF2* transcripts from promoter P1 are always derived from both parental alleles, whereas transcripts from promoters P2, P3 and P4 are always from one parental allele. These findings demonstrate that imprinting and a lack of imprinting can both occur within a single gene in a single tissue, suggesting that regional imprinting factors may be important.

We studied the allelic expression of *IGF2* transcripts derived from each of the four promoters P1–P4 using the polymorphic *Apal* restriction site<sup>19</sup>. After screening human tissues for *Apal* heterozygosity, we made full-length cDNA with primers p6 or p7 derived from exon 9 (Fig. 1a). Amplification using promoter-specific primers (Fig. 1b), followed by a nested polymerase chain reaction (PCR) using a <sup>32</sup>P-end-labelled primer p9 resulted in four promoter-specific PCR products which were labelled at one end (P1–P4; Fig. 1c). Digestion of the PCR products with *Hinf*I and *Apal* distinguishes the *A* and *B* alleles (Fig. 1c). We verified that enzyme digestion was complete by subjecting a homologous *B/B* allele control template to identical enzymatic degradation (Fig. 1d). Amplification of genomic DNA from subjects heterozygous for *Apal* using primers p5 and p9 resulted in equal amplification of both *A* and *B* alleles (Fig. 2a; DNA). In a reconstitution experiment, cDNAs from *A/A* and *B/B* homozygous livers were mixed and then amplified using promoter-

specific primers or p5. Both the *A* and the *B* alleles from fetal and adult liver were amplified with equivalent efficiency (fetal liver shown in Fig. 2a, b). Promoter-specific PCR products were verified by identifying their promoter-specific exons using multiplex PCR (data not shown) or by ascertaining their specific sizes as in the case of fetal liver from subject 1 (Fig. 2c). These PCR transcripts included the polymorphic *Apal* restriction site.

The fetal liver full-length PCR transcripts from P1 (Fig. 3a; P1, lane 1) were digested to a 141-base end-labelled fragment with *Hinf*I (lane 2) and showed that alleles *A* and *B* were both present after further digestion with *Apal* (lane 3). We did the same experiment in the presence of the internal control template depicted in Fig. 1d in order to verify that enzymatic digestion was complete (Fig. 3a, P1, lanes 4–6). In contrast, full-length PCR transcripts from P2, P3 and P4 (Fig. 3a, b, lane 1) revealed only allele *B* after complete digestion with *Hinf*I and *Apal* (lanes 3 and 6). Fetal livers from subjects 2 and 3 showed biallelic expression in transcripts from P1 but only one allele in transcripts from P3 or P4 (Table 1).

In adult liver of subject 4, P1 PCR transcripts (Fig. 3c; P1, lanes 1 and 4) showed that both alleles were present after digestion with both *Hinf*I and *Apal* (lanes 3 and 6). PCR transcripts from P3 and P4 (Fig. 3c; P3 and P4, lane 1) had only allele *A* (lanes 3 and 6). Complete digestion with *Hinf*I and *Apal* is indicated in lanes 5 and 6. P1 PCR transcripts from adult chondro-

TABLE 1 Monoallelic and biallelic expression of *A* and *B* alleles in transcripts from *IGF2* promoters P1–P4

| Subject | Tissue                  | P1 | P2 | P3 | P4 |
|---------|-------------------------|----|----|----|----|
| 1       | Fetal liver (19 weeks*) | AB | –B | –B | –B |
| 2       | Fetal liver (22 weeks*) | AB | †  | –B | †  |
| 3       | Fetal liver (21 weeks*) | AB | †  | A– | A– |
| 4       | Adult liver             | AB | †  | A– | A– |
| 5       | Chondrocyte             | AB | †  | –B | –B |
| 6       | Chondrocyte             | AB | †  | –B | –B |
| 7       | Chondrocyte             | AB | †  | A– | †  |
| 8       | Chondrocyte             | AB | †  | –B | –B |
| 9       | Chondrocyte             | AB | †  | A– | A– |
| 10      | Chondrocyte             | AB | †  | A– | †  |

Promoter-specific transcripts (P1, P2, P3 and P4) from 10 unrelated subjects who were heterozygous at the *Apal* restriction site polymorphism were analysed for the presence of *A* (undigested) and *B* (*Apal*-digested) alleles. Both alleles were present in transcripts from promoter P1 but only one allele was detected in transcripts from promoters P3 and P4. Transcripts from P2 are present in exceedingly low abundance in most tissues and could be assessed in only one case of fetal liver.

\* Weeks of gestation.

† Not detected, or abundance of transcript was too low to allow full-length amplification.

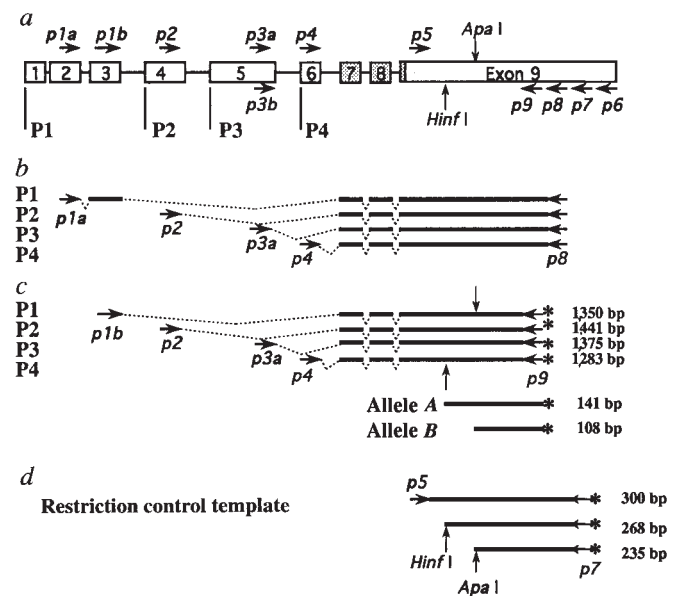
rocytes from subject 5 also showed both alleles (Fig. 3d; P1 lanes 1–6) whereas promoter P3 transcripts had only allele *B* (P3, lanes 4–6).

Table 1 summarizes the expression of *A* and *B* alleles in transcripts from each of the *IGF2* promoters in ten unrelated subjects. Although we observed nearly equal relative intensities of the *A* and *B* alleles derived from the P1 transcripts in most samples (for example, subject 5; Fig. 3d), other specimens show an unequal ratio of *A* and *B* alleles (see Fig. 3a for example).

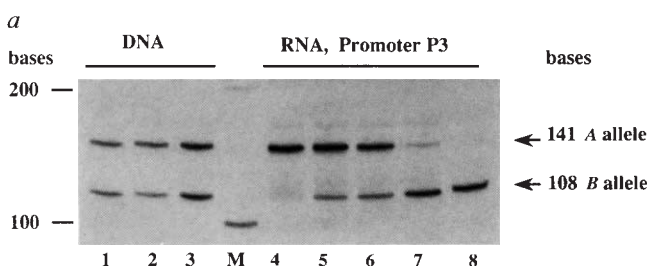
**FIG. 1** Map of *IGF2* primers. *a*, *IGF2* comprises 9 exons and 8 introns, uses four promoters (P1–P4), and encodes prepro-IGF-II (shaded exons). Parental alleles were distinguished by digestion using a common *HinfI* site and the polymorphic *ApaI* site. *b*, RNAs were reverse-transcribed using primer p6 or p7. Each promoter-specific transcript was amplified using a specific primer (p1a, p2, p3a and p4) and p8. *c*, PCR products were labelled (asterisks) separately by nested PCR using each of four primers (p1b, p2, p3a or p3b, p4) and <sup>32</sup>P-end-labelled p9. *HinfI* digestion resulted in a 141-bp end-labelled fragment which was digested further with *ApaI*. Transcripts that do not demonstrate genomic imprinting show two alleles: an *ApaI*-undigested allele, *A*, of (141 bp), and an *ApaI*-digested allele, *B*, of 108 bp. *d*, To rule out partial digestion by *ApaI*, a homologous larger control template was added before the digestion. Amplification of *B/B* fetal liver using p5 and <sup>32</sup>P-end-labelled p7 resulted in a 300-bp fragment. Complete digestion by *ApaI* yielded a single 235 bp fragment.

**METHODS.** Poly(A)<sup>+</sup> RNA was reverse-transcribed using p6 or p7 at 37 °C for 30 min followed by 5 thermal cyclings (50 °C for 20 s and 37 °C for 5 min). Aliquots of 10-fold-diluted cDNA were amplified using p8 and each of the four specific primers for 30 cycles (94 °C for 30 s, 60 °C for 30 s, and 72 °C for 90 s, followed by 72 °C for 5 min). Nested PCR (20 cycles) was performed using <sup>32</sup>P-end-labelled p9 and each of the following: p1b, p2, p3a or p3b and p4. Labelled PCR products were purified on 1% low-melting agarose gel, digested with *HinfI* or *HinfI*+*ApaI* (1 unit each) at 37 °C for 30 min and analysed on a 5% polyacrylamide-urea gel. Primers sequences: p1a: 5'-CGAATTCTGGGC-ACCAGTGACTCCCG-3'; p1b: 5'-CAGTCCTGAGGTGAGCTGCTG-3'; p2: 5'-ACCGGGCATTGCCCCAGT(TC)TCC-3'; p3a: 5'-CGTCGCACATTCGGC-(TC)(C)TCCGACT-3'; p3b: 5'-GGTTTGCGACACGACAGGGAG(AG)TG-3'; p4: 5'-TCCTCCTC(TC)TCC(AT)GCCCCAGCG-3'; p5: 5'-CTTGAGACT-

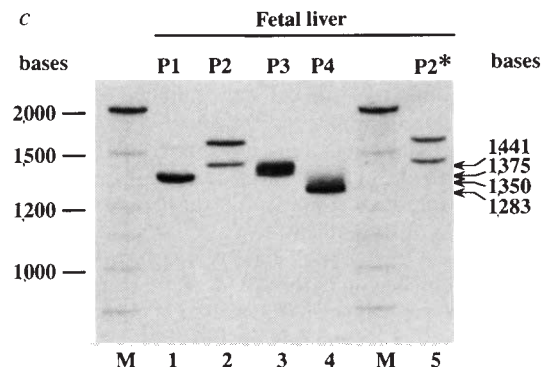
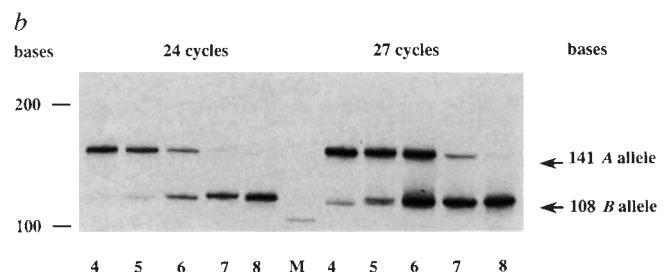
We cannot rule out the existence of factors that might differentially affect the steady-state levels of the two alleles before or during PCR amplification, nor can we exclude, in some cases, the possibility that promoter P1 is partially imprinted. It is clear, however, that in every tissue examined, all of the transcripts derived from promoter P1, regardless of their abundance, were expressed biallelically, indicating the lack of genomic imprinting. Transcripts derived from promoters P2, P3 and P4, on the other hand, were all monoallelically expressed.



T(GT)AGTCAAATTGGC-3'; p6: 5'-GACCTGTGTTTCATGTATGTGCTGCACA-3'; p7: 5'-GGGGTGAGGGT(CT)GTGCCAATTA-3'; p8: 5'-GTGATGAAAAGGGA-GTGAGGAGCC-3'; p9: 5'-GAGGAGCCAGTCTGGGTTGTTGCTA-3'.



**FIG. 2** Equal amplification of *IGF2* *A* and *B* alleles. *a*, DNAs from subjects 1, 2 and 3 (Table 1) amplified by p5 and p9 show equal intensity of *A* and *B* alleles (lanes 1–3). cDNAs from *A/A* and *B/B* homozygous fetal livers were mixed in the following ratios (*A*:*B*): 16:1, lane 4; 4:1, lane 5; 1:1, lane 6; 1:4, lane 7; and 1:16, lane 8. Lane M, Size markers. The cDNA mixtures were amplified as depicted in Fig. 1a–c using p3a, p8 and p9. Both *A* and *B* alleles from P3-derived transcripts were amplified with equivalent efficiency (lanes 4–8). Similar results were seen in P1- and P4-derived transcripts. *b*, Amplification of the cDNA mixtures (as in lanes 4–8 in *a*) using p5 and p9 also show equivalent amplification of the *A* and *B* alleles. The 1:1 mixture showed an *A*:*B* ratio of 52:48 (24 PCR cycles) and 43:57 (27 PCR cycles). The ratio of the measured intensities for the other mixtures differed from expected by an average of <5% at 24 cycles and <8% at 27 cycles. *c*, Full-length amplification of *IGF2* transcripts in fetal liver. The labelled PCR products in Fig. 1c were sized on a 3% polyacrylamide-urea gel with a <sup>32</sup>P-end-labelled 100-base ladder (BRL, Lane M). Amplified transcripts from promoter P1 show the expected 1,350 base DNA (lane 1). Transcripts from promoter P2 show the expected 1,441-base band (lane 2) and a ~1,600-base band, presumably due to alternative RNA splicing. P3 and P4 products show expected bands at 1,375 bases and 1,283 bases, respectively (lanes 3 and 4). Consistent results were also



obtained by amplification using p7 instead of p8. A representative result using p7 for promoter P2 shows both 1,441 base and ~1,600 base bands (P2\*, lane 5).



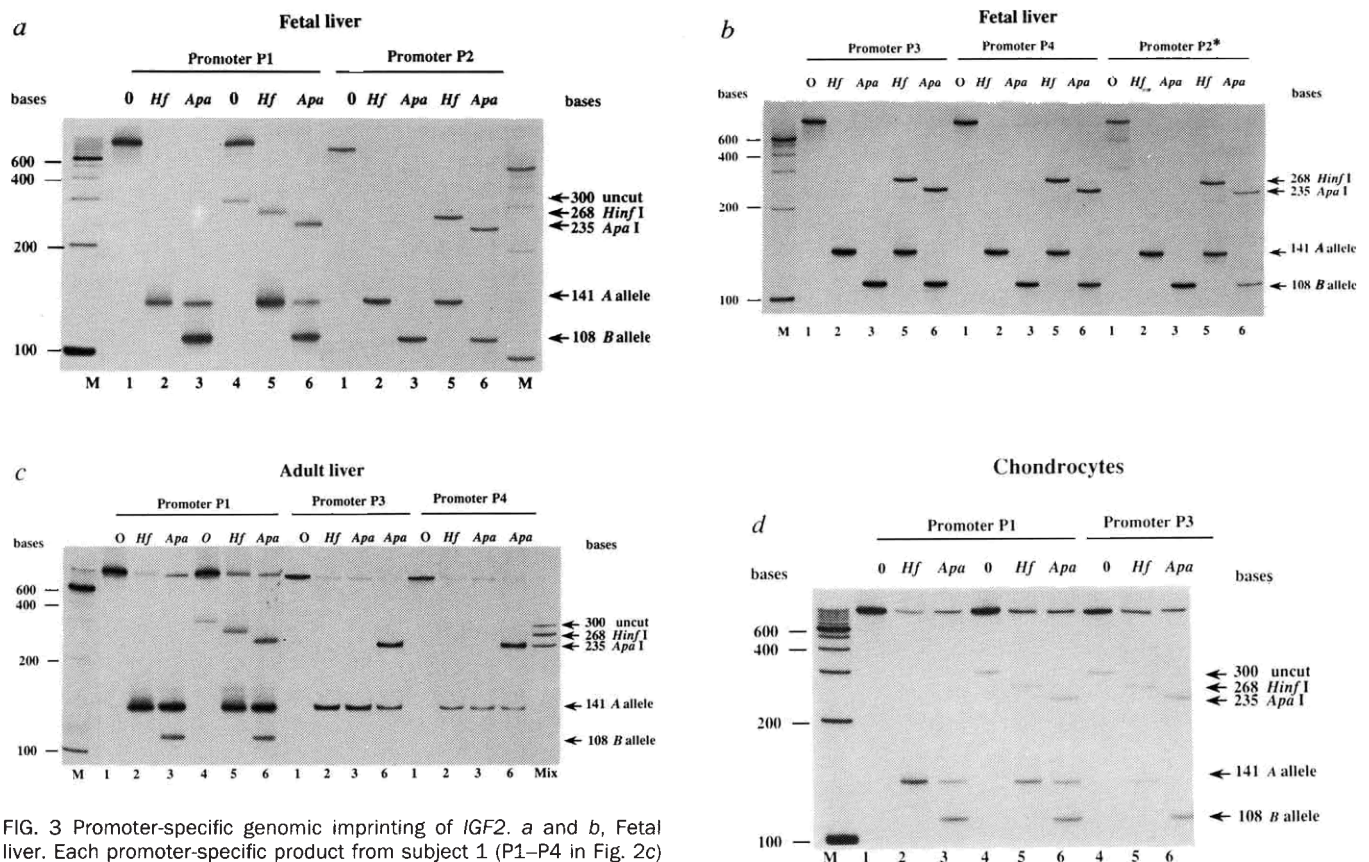


FIG. 3 Promoter-specific genomic imprinting of *IGF2*. *a* and *b*, Fetal liver. Each promoter-specific product from subject 1 (P1–P4 in Fig. 2c) was analysed to determine whether transcripts were derived from one or both parental alleles. PCR products (0 (lane 1), no enzyme) were digested with *HinfI* (lane 2) or *HinfI* + *ApaI* (lane 3). Identical digestion conditions in the presence of a 300-base internal control template (no enzyme, lane 4; *HinfI*, lane 5; *HinfI* + *ApaI*, lane 6) showed that digestion was complete. Both alleles are observed in transcripts from promoter 1 (P1, lanes 3 and 6), but transcripts from promoters P2–P4 are exclusively from allele B (P2–P4; lanes 1–6). To confirm the imprinting of promoter P2, which is of very low abundance, we also analysed the P2\* PCR product from Fig. 2c. *c*, Adult liver. Each promoter-specific PCR product from subject 4 (no enzyme, lane 1) was digested with *HinfI* (lane 2) or *HinfI* + *ApaI* (lane 3). The internal control templates in lanes 4, 5 and 6 are completely digested by *HinfI* + *ApaI*. Transcripts from

purified P1 were derived from both alleles (P1; lanes 3 and 6). Transcripts from P3 and P4 were exclusively from allele A (P3, lanes 3 and 6; P4, lanes 3 and 6). Transcripts from P2 could not be amplified from adult liver. *d*, Chondrocytes. Purified P1 and P3 promoter-specific PCR products from subject 5 (P1 and P3, lanes 1 and 4; no enzyme) show a band at 141 bases by *HinfI* digestion (P1 and P3, lanes 2 and 5). Further digestion with *ApaI* in the absence (lane 3) and in the presence of the internal control template (lane 6) allows the differentiation of the two alleles. Transcripts from P1 were from both alleles (P1, lanes 3 and 6); transcripts from P3 were exclusively from allele B (P3, lane 6). Lane M shows a 100-base ladder.

It has been shown that the maternal *IGF2* allele is normally imprinted and that the paternal allele is always transcribed<sup>15</sup>. Partial imprinting may occur when transcriptional repression of the imprinted allele is not complete, or when both alleles are expressed in some cells in the tissue while the gene is fully imprinted in other cells. Our data suggest that the partial genomic imprinting of *IGF2* reported in both normal and neoplastic tissues could also be explained by the fact that *IGF2* is expressed from multiple promoters within a single cell or tissue, and that whereas all four promoters transcribe the paternal allele, promoter usage strictly determines the transcriptional fate of the maternal allele. When P1 is used, both the maternal and paternal alleles are transcribed, but only the paternal allele is expressed from the other promoters. Choroid plexus/leptomeninges also use P1 and express *IGF2* biallelically<sup>18</sup>.

As the three *IGF2* promoters P2–P4 are clustered in a ~5-kilobase (kb) DNA region<sup>20</sup> and P1 is ~20 kb upstream, the imprinting signal(s) may be localized in a region between promoters P1 and P2. Although there is no functional mouse homologue of human P1 (ref. 21), differences in parental allelic methylation have been observed in a region ~3 kb upstream of the first mouse promoter<sup>22</sup>, which is equivalent to human promoter P2. The methylation status of this region in the human gene has not been fully investigated.

*IGF2* is imprinted in all mouse tissues except for the choroid plexus and leptomeninges<sup>23</sup>; the mechanism for biallelic expression in mice is unknown. Transcription from the maternal allele is not completely repressed in the mouse, and it has been estimated that maternal *IGF2* transcripts represent ~3% of the paternal transcripts<sup>22</sup>. In the mouse and rat, which do not have promoter P1, *IGF2* expression declines soon after birth, but species that use P1 have persistent *IGF2* expression throughout life. Most of this IGF-II is thought to originate from liver, which uses predominantly P1. The biallelic expression of *IGF2* from liver may permit more abundant synthesis of the IGF-II peptide for export into the circulation.

Genomic imprinting has previously been considered as a mechanism whereby an entire allele or even entire regions of chromosomes are transcriptionally silenced<sup>24</sup>. Our data indicate that imprinting can be restricted to a specifically delineated region of a gene. Therefore, in some circumstances, loss of genomic imprinting may simply reflect a switch in promoter usage and not a disruption of putative imprinting signals. □

Received 8 August; accepted 16 September 1994.

1. Feinberg, A. P. *Nature Genet.* **4**, 110–113 (1993).
2. Rainier, S. et al. *Nature* **362**, 747–749 (1993).
3. Ogawa, O. et al. *Nature* **362**, 749–751 (1993).

4. Eksberg, R., Shen, D. R., Fei, Y. L., Song, Q. L. & Squire, J. *Nature Genet.* **5**, 143–150 (1993).
5. Ogawa, O. *et al. Nature Genet.* **5**, 408–412 (1993).
6. Suzuki, H., Ueda, R., Takahashi, T. & Takahashi, T. *Nature Genet.* **6**, 332–333 (1994).
7. Chaillet, J. R., Vogt, T. F., Beier, D. R. & Leder, P. *Cell* **66**, 77–83 (1991).
8. Ferguson-Smith, A. C., Sasaki, H., Cattanaach, B. M. & Surani, M. A. *Nature* **362**, 751–755 (1993).
9. Li, E., Beard, C. & Jaenisch, R. *Nature* **366**, 362–365 (1993).
10. Stoger, R. *et al. Cell* **73**, 61–71 (1993).
11. Brandeis, M. *et al. EMBO J.* **12**, 3669–3677 (1993).
12. Schneid, H. *et al. J. med. Genet.* **30**, 353–362 (1993).
13. Van Dijk, M. A., Van Schaik, F. M. A., Bootsma, H. J., Holthuisen, P. & Sussenbach, J. S. *Molec. cell. Endocr.* **81**, 81–94 (1991).
14. Ohlsson, R. *et al. Nature Genet.* **4**, 94–97 (1993).
15. Giannoukakis, N., Deal, C., Paquette, J., Goodyer, C. G. & Polychronakos, C. *Nature Genet.* **4**, 98–101 (1993).
16. Kalscheuer, V. M., Mariman, E. C., Shepins, M. T., Rehder, H. & Ropers, H. H. *Nature Genet.* **5**, 74–78 (1993).
17. Sussenbach, J. S., Steenbergh, P. H. & Holthuisen, P. *Growth Reg.* **2**, 1–9 (1992).
18. Ohlsson, R., Hedberg, F., Holmgren, L., Walsh, C. & Ekstrom, T. J. *Development* **120**, 361–368 (1994).
19. Tadokoro, K., Fujii, H., Inoue, T. & Yamada, M. *Nucleic Acids Res.* **19**, 6967 (1991).
20. Sussenbach, J. S. *Prog. Growth Factor Res.* **1**, 33–48 (1989).
21. Rotwein, P. & Hall, L. J. *DNA Cell Biol.* **9**, 725–735 (1990).
22. Sasaki, H. *et al. Genes Dev.* **6**, 1843–1856 (1992).
23. DeChiara, T. M., Robertson, E. J. & Efstratiadis, A. *Cell* **64**, 849–859 (1991).
24. Cattanaach, B. M. & Beechey, C. V. *Development (suppl.)* 63–72 (1990).

ACKNOWLEDGEMENTS. We thank the Cooperative Human Tissue Network for liver specimens, K. Tsuchiya and R. L. Smith for chondrocytes, and I. H. Jin for technical assistance. This work was supported by an NIH grant and by the Research Service of the Department of Veterans Affairs.

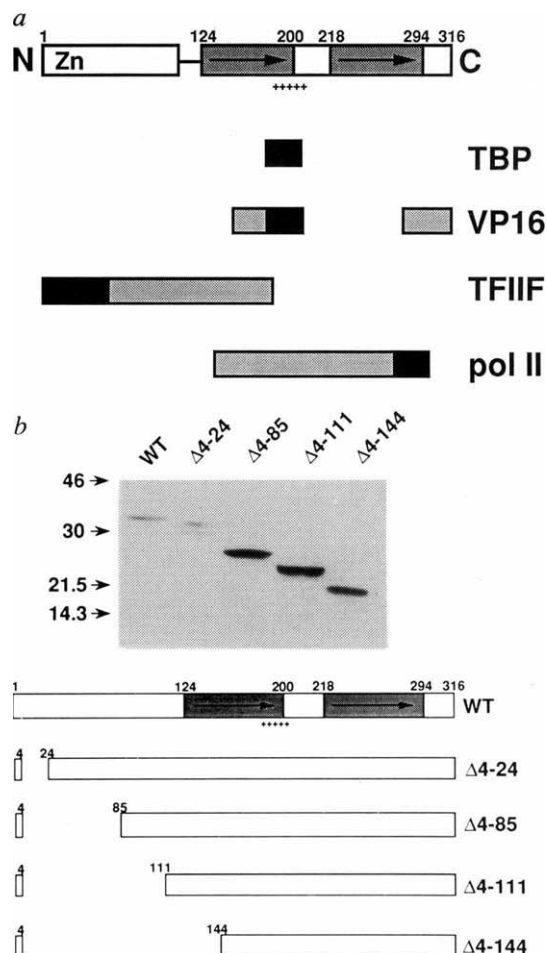
## Activator-induced conformational change in general transcription factor TFIIB

Stefan G. E. Roberts & Michael R. Green

Howard Hughes Medical Institute, Program in Molecular Medicine, University of Massachusetts Medical Center, 373 Plantation Street, Worcester, Massachusetts 01605, USA

**TRANSCRIPTIONAL activator proteins (activators) function, at least in part, by increasing preinitiation complex assembly (for reviews see refs 1–4). Previous studies have shown that an acidic activator forms a contact with the general transcription factor TFIIB (refs 5–7) and recruits it into the preinitiation complex<sup>5,8,9</sup>. Mutational studies indicate that this interaction between the acidic activator and TFIIB is required for transcriptional stimulation<sup>5–7,10,11</sup>. We show here that the acidic activator–TFIIB interaction has an additional function in preinitiation complex assembly. We provide evidence that in native TFIIB the amino- and carboxy-terminal domains are engaged in an intramolecular interaction. The acidic activator disrupts this intramolecular interaction to expose binding sites for general transcription factors that enter the preinitiation complex through association with TFIIB. Thus, the acidic activator induces a conformational change in TFIIB that drives preinitiation complex assembly forward.**

Previous structural and functional studies have shown that TFIIB has two domains: a 192-amino-acid C-terminal region containing two imperfect direct repeats, and a 124-amino-acid N-terminal region which contains a putative metal-binding site<sup>12,16</sup>. We have previously mapped the binding site for the acidic activation domain of the herpes simplex virus-1 VP16 protein to within the TFIIB C-terminal region<sup>7</sup> (Fig. 1a). Figure 1b examines the role of the TFIIB N terminus in VP16 binding using a 'far western' blotting assay. Deletion of the N-terminal 24 amino acids of TFIIB ( $\Delta$ 4–24), which disrupts the putative metal-binding site<sup>12</sup>, did not affect binding of <sup>35</sup>S-labelled GAL4–VP16. Unexpectedly, however, further deletions within the TFIIB N-terminus ( $\Delta$ 4–85,  $\Delta$ 4–111 and  $\Delta$ 4–144) significantly enhanced the affinity of TFIIB for VP16. Thus, the TFIIB N terminus is not required, and actually inhibits, VP16 binding.



**FIG. 1** The TFIIB N-terminal region inhibits VP16 binding. **a**, TFIIB regions involved in transcription factor interactions. Schematic diagram of the domains in TFIIB required for interaction with TATA-binding protein (TBP), VP16, TFIIF and RNA polymerase II (pol II). Black boxes, regions in which amino-acid substitutions interfere with binding. The basic region (plus signs), direct repeats (arrows) and a potential zinc-binding site (Zn) of TFIIB are indicated. **b**, Far western blotting analysis. Bacterial lysates containing equal amounts of the indicated TFIIB derivatives were resolved by SDS–PAGE, transferred to an immobilized membrane, and probed with <sup>35</sup>S-methionine-labelled GAL4–VP16. After extensive washing the membrane was dried and autoradiographed. Relative molecular mass standards  $\times 10^{-3}$  are shown on the left. WT, wild type.

**METHODS.** Binding sites in TFIIB for the proteins indicated were compiled from data described here and elsewhere<sup>7,12</sup>. Construction and expression of TFIIB deletion mutants have been described<sup>12</sup>. Bacterial lysates containing equal amounts of the TFIIB deletion mutants were fractionated on a 10% SDS–polyacrylamide gel and transferred to the Immobilon P membrane (Millipore). The membrane was incubated in far western buffer (150 mM KCl, 20 mM HEPES pH 7.5, 5 mM MgCl<sub>2</sub>, 0.5 mM EDTA, 0.05% NP40, 1 mM dithiothreitol (DTT) and 1 mM PMSF) containing 5% BSA for 2 h at 4 °C. <sup>35</sup>S-methionine-labelled GAL4–VP16 (ref. 7) in far western buffer containing 0.5% BSA was added and incubation at 4 °C continued for 4 h. The membrane was washed twice for 15 min each in far western buffer, dried and autoradiographed.

One explanation for the results of Fig. 1b is that the N terminus of TFIIB masks the VP16-binding site through an intramolecular interaction. To test this idea, we carried out protein-affinity chromatography experiments in which the N-terminal 124 amino acids of TFIIB, or deletions therein, were fused to glutathione *S*-transferase (GST) and the various fusion proteins immobilized on glutathione–agarose beads. Figure 2a shows that a 193-residue TFIIB C-terminal domain bound to the TFIIB N-terminal domain but not the control GST protein. These data



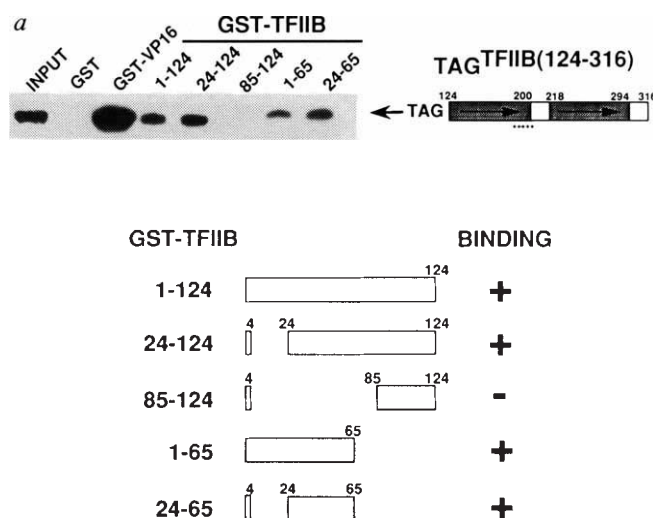


FIG. 2 Intramolecular interaction between TFIIB N- and C-terminal domains. **a**, N-terminal region. The N terminus of TFIIB (residues 1–124) and the indicated deletion mutants were expressed as GST fusion proteins and coupled to glutathione-agarose beads. Immobilized fusion proteins were incubated with the epitope-tagged C terminus of TFIIB (TAG-TFIIB (124–316)). GST and GST-VP16 were included as negative and positive controls, respectively. After extensive washing, the bound TAG-TFIIB (124–316) was eluted, resolved by SDS-PAGE and detected by immunoblotting. **b**, C-terminal region. Epitope-tagged TFIIB C terminus and the indicated deletion mutants were incubated with immobilized GST or GST-TFIIB(residues 1–124). Bound proteins were analysed as in **a**. The diagram at the bottom summarizes the intramolecular binding sites (black bars) as determined in **a** and **b**. **c**, VP16 competition assay. Epitope-tagged TFIIB C terminus (TAG-TFIIB(124–316)) was bound to immobilized GST-TFIIB(1–124) as in **a**. Increasing amounts (1  $\mu$ g, 5  $\mu$ g and 25  $\mu$ g) of GAL4 (1–93) or GAL4 (1–93)-VP16 were then added and incubation allowed to continue for an additional 15 min. After further washing, remaining bound TFIIB C terminus was eluted and detected as in **a**.

**METHODS.** The TFIIB N terminus (residues 1–124) and deletions were cloned into pGEX2T and expressed proteins purified as described<sup>5</sup>. Epitope-tagged TFIIB C terminus (TAG-TFIIB(124–316)) was constructed by cloning residues 124–316 of TFIIB in-frame to the 12-amino-acid T7 tag in pET 5a (Novagen). Previously described deletions in the TFIIB C terminus<sup>12</sup> were cloned into pET5a and used to prepare bacterial

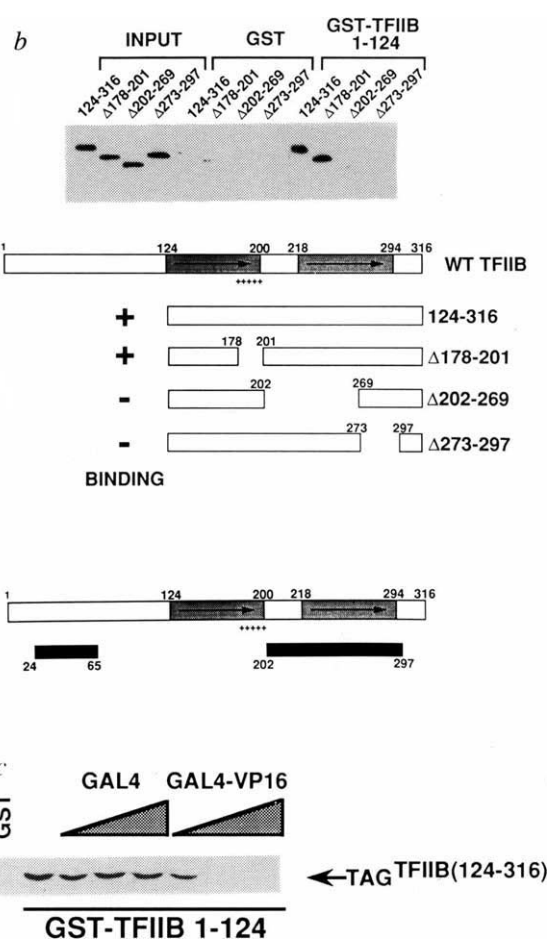
lysates. The protein binding assay was as described previously<sup>12</sup> using 2  $\mu$ g of GST fusion protein immobilized on 20  $\mu$ l glutathione-agarose beads. Tagged TFIIB C terminus was detected with an  $\alpha$ -T7 monoclonal antibody (Novagen) and visualized by enhanced chemiluminescent detection (ECL; Amersham). GAL4(1–93) and GAL4(1–93)-VP16 were prepared as histidine-tagged proteins as described<sup>24</sup>.

provide direct evidence for the proposed intramolecular interaction between the TFIIB N- and C-terminal domains. Analysis of TFIIB N-terminal derivatives (Fig. 2a) identified a 40-amino-acid N-terminal region (residues 24–65) that was sufficient for the interaction.

Figure 2b maps the portion of the TFIIB C-terminal domain required for the intramolecular interaction. Deletion of the basic region ( $\Delta$ 178–201), which is required for VP16 binding<sup>7</sup>, did not affect binding to the TFIIB N terminus. However, deletions in the second direct repeat ( $\Delta$ 202–269 and  $\Delta$ 273–297) abolished the interaction. Thus, the TFIIB N terminus interacts with the second direct repeat of the C-terminal domain (summarized in Fig. 2b, bottom).

Our findings suggested that the N terminus of TFIIB may compete with VP16 for binding to the TFIIB C-terminal domain, and the following experiment showed that this was indeed the case. In a competition experiment in which the TFIIB C-terminal domain was bound to the immobilized GST-TFIIB N terminus (1–124) and challenged with GAL4(1–93)-VP16, or the control GAL4(1–93), addition of GAL4(1–93) VP16, but not GAL4(1–93), inhibited the interaction between the TFIIB N- and C-terminal domains (Fig. 2c).

As the competition data of Fig. 2c strongly suggested that VP16 could disrupt the TFIIB intramolecular interaction, we used a partial proteolysis assay<sup>17,18</sup> to test whether this disruption



involved a conformational change in TFIIB. A TFIIB protein epitope-tagged at the C terminus was subjected to limited digestion with *Staphylococcus aureus* V8 protease in the presence of GST-VP16 or a control GST protein. Figure 3a shows the TFIIB protease cleavage pattern in the presence of increasing amounts of GST-VP16. The major proteolytic digestion products were unaffected by VP16 and provided useful internal controls. A single proteolytic cleavage product was observed which was dependent on the concentration of VP16 (VP16-induced). The size of this fragment indicates that the proteolytic cleavage site is at (or near) a glutamate residue at position 51, which is within the region of TFIIB involved in the intramolecular interaction (24–65). Thus, binding of VP16 to the TFIIB C-terminal domain induces a conformational change in the TFIIB N terminus.

To confirm that the VP16-induced proteolytic cleavage site was dependent on interaction between VP16 and TFIIB, we analysed VP16 (refs 5, 6, 10, 11) and TFIIB<sup>7</sup> mutants. A VP16 derivative containing a phenylalanine-to-proline substitution at position 442 ( $\Delta$ 456FP442) has a decreased affinity for TFIIB<sup>5,6</sup> and cannot activate transcription<sup>10,11</sup>. When a control GST-VP16 $\Delta$ 456 protein, or GST-VP16 $\Delta$ 456FP442, was incubated with epitope-tagged TFIIB and subjected to partial proteolysis, VP16 $\Delta$ 456, but not  $\Delta$ 456FP442, induced the protease cleavage site (Fig. 3b).

We have described TFIIB amino-acid substitution mutants with decreased affinity for VP16 that support basal, but not activated, transcription<sup>7</sup>. One such mutant, TFIIB-R185E/R193E, was epitope-tagged and tested for binding to GAL4-VP16 using far western blotting (Fig. 3c, top panel). In agreement with our previous protein affinity-chromatography data<sup>7</sup>, binding of TFIIB-R185E/R193E was significantly less

than that of wild-type TFIIB (Fig. 3c, compare lanes 1 and 2). Figure 3c (bottom panel) shows that wild-type TFIIB, but not TFIIB-R185E/R193E, gave rise to the VP16-induced protease cleavage site. The combined results of Fig. 3 indicate that the VP16-induced protease cleavage site required the VP16-TFIIB interaction.

The model (in Fig. 4) summarizes our results. In the absence

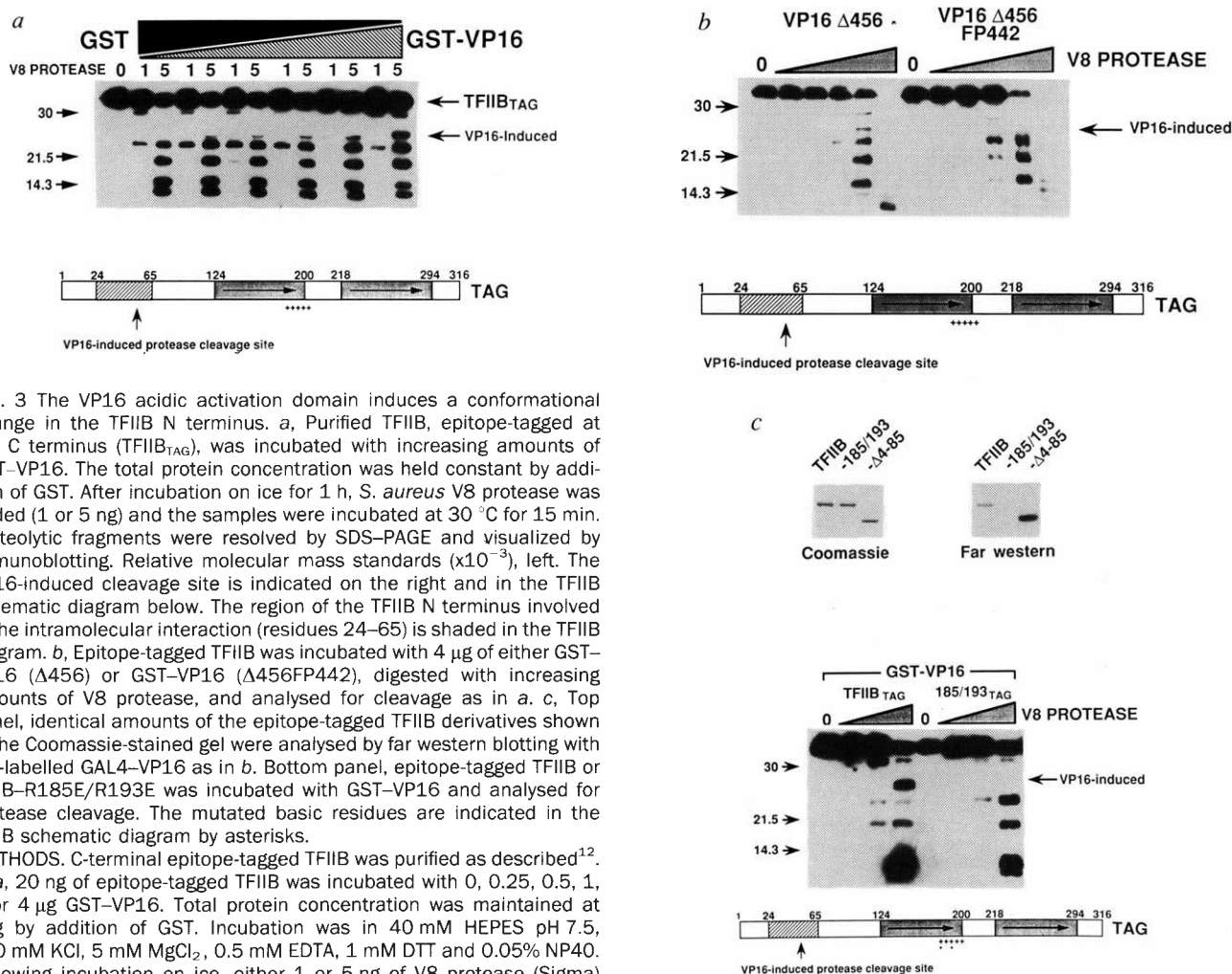


FIG. 3 The VP16 acidic activation domain induces a conformational change in the TFIIB N terminus. **a**, Purified TFIIB, epitope-tagged at the C terminus (TFIIB<sub>TAG</sub>), was incubated with increasing amounts of GST-VP16. The total protein concentration was held constant by addition of GST. After incubation on ice for 1 h, *S. aureus* V8 protease was added (1 or 5 ng) and the samples were incubated at 30 °C for 15 min. Proteolytic fragments were resolved by SDS-PAGE and visualized by immunoblotting. Relative molecular mass standards ( $\times 10^{-3}$ ), left. The VP16-induced cleavage site is indicated on the right and in the TFIIB schematic diagram below. The region of the TFIIB N terminus involved in the intramolecular interaction (residues 24–65) is shaded in the TFIIB diagram. **b**, Epitope-tagged TFIIB was incubated with 4  $\mu$ g of either GST-VP16 ( $\Delta 456$ ) or GST-VP16 ( $\Delta 456$ FP442), digested with increasing amounts of V8 protease, and analysed for cleavage as in **a**. **c**, Top panel, identical amounts of the epitope-tagged TFIIB derivatives shown in the Coomassie-stained gel were analysed by far western blotting with <sup>35</sup>S-labelled GAL4-VP16 as in **b**. Bottom panel, epitope-tagged TFIIB or TFIIB-R185E/R193E was incubated with GST-VP16 and analysed for protease cleavage. The mutated basic residues are indicated in the TFIIB schematic diagram by asterisks.

**METHODS.** C-terminal epitope-tagged TFIIB was purified as described<sup>12</sup>. In **a**, 20 ng of epitope-tagged TFIIB was incubated with 0, 0.25, 0.5, 1, 2 or 4  $\mu$ g GST-VP16. Total protein concentration was maintained at 4  $\mu$ g by addition of GST. Incubation was in 40 mM HEPES pH 7.5, 120 mM KCl, 5 mM MgCl<sub>2</sub>, 0.5 mM EDTA, 1 mM DTT and 0.05% NP40. Following incubation on ice, either 1 or 5 ng of V8 protease (Sigma) was added and the samples were incubated at 30 °C for an additional 15 min. Digestion was ended by adding SDS-PAGE loading dye followed by further incubation at 100 °C. SDS-PAGE and immuno-detection was as for Fig. 2. TFIIB R185E/R193E and  $\Delta 4-85$  have been described<sup>7</sup> and were epitope-tagged. Purification of GST-VP16, GSTVP16 $\Delta 456$  and GSTVP16 $\Delta 456$ FP442 has been described<sup>5</sup>. Protease assays were as in **a**, with 20 ng of TFIIB (or 20 ng R185E/R193E), but using 4  $\mu$ g of GST-VP16, GSTVP16 $\Delta 456$  or GSTVP16 $\Delta 456$ FP442. In **b**, 0, 10<sup>-6</sup>, 10<sup>-5</sup>, 10<sup>-4</sup>, 10<sup>-3</sup> or 10<sup>-2</sup>  $\mu$ g of V8 protease was added in each lane, and in **c**, 0, 10<sup>-5</sup>, 10<sup>-4</sup> and 10<sup>-3</sup>  $\mu$ g of V8 protease was used.

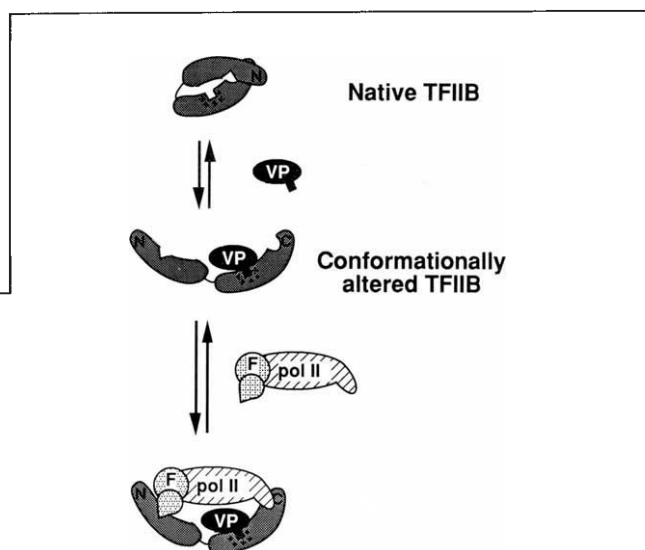


FIG. 4 Model for acidic activator-induced TFIIB conformational change. In native TFIIB, the binding sites for TFIIF and RNA polymerase II are engaged in an intramolecular interaction and thus assembly of TFIIF/pol II is limited. In the absence of an activator, only a small fraction of TFIIB is in the alternative 'active' conformation in which the TFIIF- and RNA polymerase II-binding sites are exposed. Binding of the activator alters the conformation of TFIIB to the active form, enabling TFIIF/pol II to assemble efficiently, driving preinitiation complex assembly forward.



of the activator, TFIIB adopts a conformation that inefficiently supports further assembly of the preinitiation complex. In this conformation the binding sites for TFIIF (N-terminal) and RNA polymerase II (C-terminal<sup>12</sup>) interact with each other, blocking access to TFIIF-RNA polymerase II.

An acidic activator binds directly to TFIIB, disrupts the intramolecular interaction and exposes the binding sites for TFIIF and RNA polymerase II. Thus, an acidic activator affects TFIIB assembly, both quantitatively, by recruiting TFIIB into the preinitiation complex<sup>5,8,9</sup>, and qualitatively, by altering TFIIB conformation in a manner that drives preinitiation complex assembly forward. This latter function of the activator-TFIIB interaction explains why raising the concentration of TFIIB *in vitro*<sup>8,19,20</sup>, or overexpressing TFIIB *in vivo*<sup>21</sup>, does not overcome the requirement for an activator.

Activators function during at least two steps of preinitiation complex assembly<sup>8,9,22</sup>, first to recruit TFIIB to the promoter and second to recruit GTFs that assemble after TFIIB. Interestingly, the TBP-associated factors (TAFs) participate in this second step. Thus, one function of TAFs may be to facilitate the activator-induced conformational change of TFIIB, perhaps by a direct TAF-TFIIB interaction. Consistent with this possibility, it has been shown that *Drosophila* TAF<sub>II</sub>40 interacts directly with TFIIB<sup>23</sup>. Future studies using purified preinitiation complexes<sup>5,8</sup> will test this idea and other aspects of our model. □

Received 14 June; accepted 26 August 1994.

- Kingston, R. E. & Green, M. R. *Curr. Biol.* **4**, 325-332 (1994).
- Hori, R. & Carey, M. *Curr. Opin. Genet. Dev.* **4**, 236-244 (1994).
- Zawel, L. & Reinberg, D. *Prog. Nucleic Acids Res. molec. Biol.* **44**, 67-108 (1993).
- Tjian, R. & Maniatis, T. *Cell* **77**, 5-8 (1994).
- Lin, Y.-S. & Green, M. R. *Cell* **64**, 971-981 (1991).
- Lin, Y.-S., Ha, I., Maldonado, E., Reinberg, D. & Green, M. R. *Nature* **353**, 569-571 (1991).
- Roberts, S. G. E., Ha, I., Maldonado, E., Reinberg, D. & Green, M. R. *Nature* **363**, 741-744 (1993).
- Choy, B. & Green, M. R. *Nature* **366**, 531-536 (1993).
- Kim, T. K. et al. *Nature* **369**, 252-255 (1994).
- Cress, W. D. & Triezenberg, S. J. *Science* **251**, 87-90 (1990).
- Jiang, Y., Triezenberg, S. L. & Gralla, J. D. *J. Biol. Chem.* **269**, 5505-5508 (1994).
- Ha, I. et al. *Genes Dev.* **7**, 1021-1032 (1993).
- Barberis, A., Muller, C. W., Harrison, S. C. & Ptashne, M. *Proc. natn. Acad. Sci. U.S.A.* **90**, 5628-5632 (1993).
- Buratowski, S. & Zhou, H. *Proc. natn. Acad. Sci. U.S.A.* **90**, 5633-5637 (1993).
- Hisatake, K., Roeder, R. G. & Horikoshi, M. *Nature* **363**, 744-747 (1993).
- Malik, S., Lee, D. K. & Roeder, R. G. *Molec. cell. Biol.* **13**, 6253-6259 (1993).
- Tan, S. & Richmond, T. J. *Cell* **62**, 367-377 (1990).
- Shuman, J. D., Vinson, C. R. & McKnight, S. L. *Science* **249**, 771-774 (1990).
- Wampler, S. L. & Kadonaga, J. T. *Genes Dev.* **6**, 1542-1552 (1992).
- White, J. et al. *EMBO J.* **11**, 2229-2240 (1992).
- Colgan, J., Wampler, S. & Manley, J. L. *Nature* **362**, 549-553 (1993).
- Kim, T. K. & Roeder, R. G. *Proc. natn. Acad. Sci. U.S.A.* **91**, 4170-4174 (1994).
- Goodrich, J. A., Hoey, T., Thut, C. J., Admon, A. & Tjian, R. *Cell* **75**, 519-530 (1993).
- Reece, R. J., Rickles, R. J. & Ptashne, M. *Gene* **126**, 105-107 (1993).

**ACKNOWLEDGEMENTS.** We thank B. Choy and J. Reese for helpful suggestions throughout this work, members of M.R.G.'s laboratory for comments on the manuscript and T. O'Toole for secretarial assistance. S.G.E.R. is the recipient of a senior postdoctoral fellowship from the Massachusetts Division of the American Cancer Society. This work was supported by grants from the NIH to M.R.G.

## ERRATUM

### Crystal structure of the extracellular region of human tissue factor

**K. Harlos, D. M. A. Martin, D. P. O'Brien, E. Y. Jones, D. I. Stuart, I. Polikarpov, A. Miller, E. G. D. Tuddenham & C. W. G. Boys**

*Nature* **370**, 662-666 (1994)

The last line of the Acknowledgements section was omitted from this Letter. This read: "C. W. G. Boys was supported by the Wellcome Trust." □

## GUIDE TO AUTHORS

PLEASE follow these guidelines so that your manuscript may be handled expeditiously.

*Nature* is an international journal covering all the sciences. Contributors should therefore bear in mind those readers who work in other fields and those for whom English is a second language, and write clearly and simply, avoiding unnecessary technical terminology. Space in the journal is limited, making competition for publication severe. Brevity is highly valued. One printed page of *Nature*, without display items, contains about 1,300 words.

Manuscripts are selected for publication according to editorial assessment of their suitability and reports from independent referees. They can be sent to London or Washington and should be addressed to the Editor. Manuscripts may be dealt with in either office, depending on the subject matter, and will where necessary be sent between offices by overnight courier. High priority cannot be given to presubmission enquiries; in urgent cases they can be made in the form of a one-page fax. All manuscripts are acknowledged on receipt but fewer than half are sent for review. Those that are not reviewed are returned as rapidly as possible so that they may be submitted elsewhere without delay. Contributors may suggest reviewers; limited requests for the exclusion of specific reviewers are usually heeded. Manuscripts are usually sent to two or three reviewers, chosen for their expertise rather than their geographical location. Manuscripts accepted for publication are processed from the London office.

*Nature* requests authors to deposit sequence and crystallographic data in the databases that exist for this purpose.

Once a manuscript is accepted for publication, contributors will receive proofs in about 4 weeks. *Nature's* staff will edit manuscripts with a view to brevity and clarity, so contributors should check proofs carefully. Manuscripts are generally published 2-3 weeks after receipt of corrected proofs. *Nature* does not exact page charges. Contributors receive a reprint order form with their proofs; reprint orders are processed after the manuscript is published and payment received.

### Categories of paper

**Review Articles** survey recent developments in a field. Most are commissioned but suggestions are welcome in the form of a one-page synopsis addressed to the Reviews Coordinator. Length is negotiable in advance.

**Progress** articles review particularly topical developments for a nonspecialist readership. They do not exceed 4 pages in length. Suggestions may be made to the Reviews Coordinator in the form of a brief synopsis.

**Articles** are research reports whose conclusions are of general interest and which are sufficiently rounded to be a substantial advance in understanding. They should not have more than 3,000 words of text (not including figure legends) or more than six display items and should not occupy more than five pages of *Nature*.

Articles start with a heading of 50-80 words written to advertise their content in general terms, to which editors will pay particular attention. The heading does not usually contain numbers, abbreviations or measurements. The introduction to the study is contained in the first two or three paragraphs of the article, which also briefly summarizes its results and implications. Articles have fewer than 50 references and may contain a few short subheadings.

**Letters to Nature** are short reports of outstanding novel findings whose implications are general and important enough to be of interest to those outside the field. Letters should have 1,000 or fewer words of text and four or fewer display items. The first paragraph describes, in not more than 150 words and without the use of abbreviations, the background, rationale and chief conclusions of the study for the particular benefit of non-specialist readers. Letters do not have subheadings and contain fewer than 30 references.

**Commentary** articles deal with issues in, or arising from, research that are also of interest to readers outside research.

**News and Views** articles inform nonspecialist readers about new scientific advances, sometimes in the form of a conference report. Most are commissioned but proposals can be made in advance to the News and Views Editor.

**Scientific Correspondence** is for discussion of topical scientific matters, including those published in *Nature*, and for miscellaneous contributions. Priority is given to letters of fewer than 500 words.

### Preparation of manuscripts

All manuscripts should be typed, double-spaced, on one side of the paper only. An original and four copies are required, each accompanied by artwork. If photographs are included, five sets of originals are required; for line drawings, one set of originals and four good-quality photocopies are acceptable. Reference lists, figure legends and tables should all be on separate sheets, all of which should be double-spaced and numbered. Three copies of relevant manuscripts in press or submitted for publication elsewhere should be included with submitted manuscripts, clearly marked as such. Five copies of revised and resubmitted manuscripts, labelled with their manuscript numbers, are required, together with five copies of a letter detailing the changes made.

**Titles** are brief and simple. Active verbs, numerical values, abbreviations and punctuation are to be avoided. Titles should contain one or two key words for indexing purposes.

**Artwork** should be marked individually and clearly with the author's name and, when known, the manuscript number. Ideally, no figure should be larger than 28 by 22 cm. Figures with several parts are to be avoided and are permitted only if the parts are closely related, either experimentally or logically. Unlettered originals of photographs should be provided. Suggestions for cover illustrations, with captions and labelled with the manuscript number, are welcome. Original artwork is returned when a manuscript cannot be published.

Protein/nucleotide sequences should ideally be in the three-letter and not the single-letter code for amino acids. One column width of *Nature* can accommodate 20 amino acids or 60 base pairs.

**Colour artwork.** A charge of £500 per page is made as a contribution towards the cost of reproducing colour figures. Inability to pay these costs will not prevent publication of essential colour figures if the circumstances are explained. Proofs of colour artwork may be sent to contributors under separate cover from their galley proofs.

**Figure legends** should not exceed 300 words and ideally should be shorter. The figure is described first, then, briefly, the method. Reference to a method published elsewhere is preferable to a full description. Methods are not described in the text.

**References** are numbered sequentially as they appear in the text, followed by those in tables and finally by those in figure legends. Only papers published or in the press are numbered and included in the reference list. All other forms of reference should be cited in the text as a personal communication, manuscript submitted or in preparation. Text is not included in reference lists. References are abbreviated according to the *World List of Scientific Periodicals* (Butterworths, London, 1963-65). The first and last page numbers are included; reference to books includes publisher, place and date.

**Abbreviations**, symbols, units and Greek letters should be identified the first time they are used. Acronyms should be avoided whenever possible and, if used, defined. Footnotes are not used in the text.

**Acknowledgements** are brief and appear after the reference list; grant and contribution numbers are not allowed.

**Supplementary information** is material relevant to Articles or Letters which cannot, for lack of space, be published in full, but which is available from *Nature* on request.

**Submission.** Manuscripts can be sent to the Editor at 4 Little Essex Street, London WC2R 3LF, UK or at 1234 National Press Building, Washington, DC 20045, USA. A telephone and fax number should be included. Manuscripts or proofs sent by air courier to London should be declared as 'manuscripts' and 'value \$5' to prevent the imposition of import duty and value-added tax.

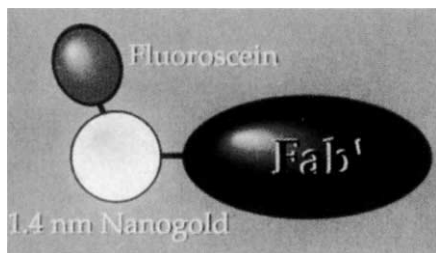
# New product listings

Recent product launches include an immunoprobe that combines fluorescence and gold into a single immunoprobe, a selection of software programs and a real-time, 3-D microscope.

BIOCONTROL is a new **rechargeable multichannel pipette** designed by Finnpiptette and available through Life Sciences International (*Reader Service No. 100*). A dual-stop trigger mechanism is said to use natural hand movement to increase comfort and cut down on the likelihood of repetitive stress injuries. It can be set to regular or stepper modes, with three pipetting speeds for either forward or reverse pipetting. Two volume ranges are available: the Biocontrol 50 (5–50 µl) and the 300 (50–300 µl). The pipette can be recharged in about an hour using the rapid recharge feature; overnight recharge is activated automatically when the pipette is placed on the stand.

## Tools for the immunologist

FluoroNanogold is a new **immunoprobe** from Nanoprobes that combines a fluorophore and a gold particle in a single antibody probe, so that two detection modes are available from one labelling experiment (*Reader Service No. 101*). Consequently, a sample



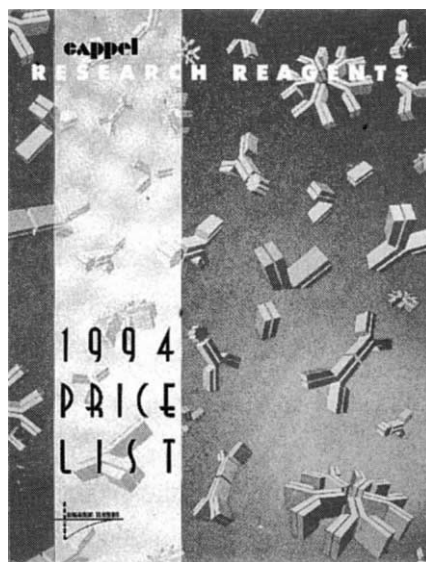
FluoroNanogold

may be analysed by fluorescence microscopy in order to show cellular antigen distribution, then by electron microscopy for high-resolution ultrastructural analysis. Nanoprobes says that the Nanogold particle does not quench the fluorescence, and the fluorophore, Nanogold particle and antibody are covalently attached for stability and long shelf life. Fab' fragments are used to maintain a small overall probe size, and are said to show good penetration of tissue. Currently available as goat anti-mouse, goat anti-rabbit, rabbit anti-goat and streptavidin conjugates with fluorescein.

**Anti-thymine dimer monoclonal antibody** is now available unlabelled or labelled with horseradish peroxidase from Kamiya Biomedical (*Reader Service No. 102*). Use of the monoclonal represents a new way to detect DNA probes in *in situ* hybridization, Southern/northern hybridization, dot blotting, DNA sequencing and chromosome mapping. For this method, the DNA probe is thymine–thymine dimerized under ordinary ultraviolet light instead of directly labelled

by the usual, time-consuming procedures. After hybridization using standard methods, the TT-dimerized probes are detected by the anti-thymine dimer monoclonal antibody.

A new line of **primary antibodies** to intermediate filaments, oncogene proteins, nuclear antigens, contractile proteins and structural proteins is available from Organon Teknika-Cappel. These are among the new products featured in the company's latest



Immunochemicals: primary antibodies.

catalogue update (*Reader Service No. 103*). These products are designed for use in immunoblotting and immunohistochemistry applications. Some can also be used in ELISAs and/or flow cytometry. The catalogue update/price list also includes new murine antibody reagents, apo/lipoproteins and other monoclonal antibodies.

Harlan Bioproducts for Science has an **inhibin assay** designed specifically to detect dimeric (bioactive) human inhibin in serum or plasma. (*Reader Service No. 104*). Historically, radioimmunoassay has been used, but the polyclonal antibody used has not been able to distinguish between the dimeric form of the molecule and the large excess of free alpha subunit forms found in body fluids. Harlan says that its antibodies make this distinction possible.

DLD Diagnostika offers a <sup>125</sup>I radio-receptor assay for the **quantitative determination of acetylcholine receptor auto-antibodies** in serum and plasma (*Reader Service No. 105*). Detergent-solubilized re-

ceptors from a human cell line are labelled with <sup>125</sup>I alpha-bungarotoxin (a snake venom). Five microlitres patient serum are incubated with the labelled receptor. The resulting complex of labelled receptor and the receptor antibody are immunoprecipitated with anti-human IgG — the radioactivity of the precipitate being directly proportional to the concentration of antibodies. The procedure takes 4 h.

Two novel **inositol polyphosphates** — [2–3H(N)] IP7 and [2–3H(N)] IP8 are now available from Du Pont and should be of interest to researchers carrying out intracellular protein binding studies to investigate the regulation of protein trafficking (*Reader Service No. 106*). Both compounds are thought to be important factors that serve in the regulation of cell function, either as cellular signalling molecules or as potential energy sources in specific cell cycles. [2–3H(N)] IP7 binds the golgi protein coatomer and modulates coatomer activity on K<sup>+</sup>-selective channels.

CD11b (C3b receptor) is a member of a family of functionally important leucocyte integrins that share a common β-subunit and participate in cellular adhesion. A new **anti-CD11b antibody** is now available from Cymbus Bioscience which has been used in indirect immunofluorescence, immunohistological staining techniques and is supplied conjugated to FITC or R-PE for flow cytometry (*Reader Service No. 107*).

## Software

With Jandel Scientific's SigmaScan/Image **image measurement software** for Microsoft Windows users can automatically measure digitized images, charts and graphs (*Reader Service No. 108*). Version 1.2 has over a dozen new features including grey-level-based 'floodfill' measurements, multiple image splicing, the Jandel Application Manager, the ability to save colour and grey

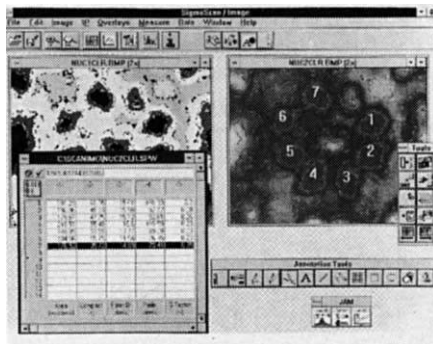
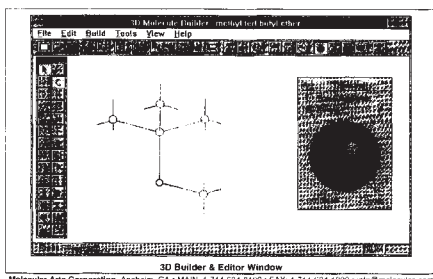
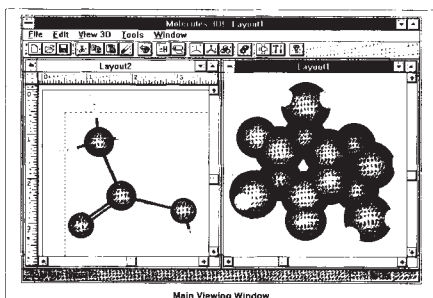


Image measurement software.



level data for each pixel in an image to and from the data worksheet and enhancements to the program's image annotation capabilities. The program comes with a built-in 65,000 row by 16,000 column worksheet with transform language and a data plotting feature. Measurements are automatically collected in the worksheet for subsequent analysis or plotting. Over 50 mathematical transforms are available for data analysis. SigmaScan/Image accepts colour and monochrome images in TIFF, BMP, TGA, PCX and GIF image file formats. This could be useful when working with graphs, electrophoresis gels, charts and photographs saved as digital files from instruments, scanners, disk or CD-ROM.

The new version of Molecules-3D, Molecular Arts Corporation's **three-dimensional molecular model building software** for Microsoft Windows, offers users over 30 new features and upgraded capabilities and is designed for scientists, engineers, academics and students in the chemistry field (*Reader Service No. 109*). With Molecules-3D, users can create Lewis structure diagrams; build, display, render and manipulate three-dimensional (3-D) structures; create graphic illustrations that are of a publication quality; export (and import) graphic images to (and from) third-party software applications; and produce



### 3-D molecular model building package.

page layouts and materials for presentation. This latest version, 2.0, offers an expandable Structure Library of over one hundred 3-D molecular structures of common, scientific or topical interest, and one-button wireframe, ball-and-stick and spacefill rendering. Other tools include 3D Builder, a Molecular Syntax Checker that identifies potential structural problems, the Visual 3D Editor for the evaluation and interactive adjustment of molecular geometries and 3D Cleanup which provides an evaluation of

molecular geometries and forces, automatically adjusting atom positions to provide more accurate conformations. Version 2.0 works on IBM or compatible PCs that support Microsoft Windows.

A new generation of **computer-aided chemistry software** based on Apple Power Macintosh computers is now available from CAChe Scientific (*Reader Service No. 110*). Called the Power Mac WorkSystem, the new system is designed to bring the company's intuitive user interface, stereo three-dimensional display and suite of applications for molecular modelling and prediction of chemical properties, structure and



**Chemistry WorkSystem** — see CAChe.

reactivity to the chemists work area. CAChe ProjectLeader, the company's property-driven experiment manager is also included in the release for PowerMac. The single-user version of CAChe's WorkSystem software, including ProjectLeader is priced at US\$29,990; special pricing is available for academic institutions.

HyperNMR is a **software package for the prediction of one-dimensional spectra** that is available from Phase Separations, and which can be used as a standalone program or in conjunction with HyperChem (*Reader Service No. 111*). NMR spectra can be simulated by HyperNMR based on the computed chemical shifts and spin-coupling constants or data obtained from elsewhere. There is no fixed limit on the number of nuclei in the NMR simulation: the maximum practical number of nuclei is said to depend only on computer memory. HyperNMR allows the user to display an NMR spectrum as individual peaks, as an envelope, or as peaks plus envelope; choose a gaussian or lorentzian distribution for line widths; select and deselect peaks interactively; zoom and pan; and copy and paste spectra.

Lab Assistant from PDMS is a Microsoft Windows-based **program designed specifically for dealing with laboratory and field data** (*Reader Service No. 112*). It provides an integrated solution to the storage of data and routine calculation procedures. The program consists of a base module, the 'engine' for the product, and a variety of 'work' modules which carry out specific calcula-

tion tasks. Being modular means that the program can be applied to most kinds of data by the design of specific modules to meet specific needs without the need for a new application. Lab Assistant sells for UK£149, although for a limited period it will be available at an introductory price of £99.

### ■ In a spin

Eppendorf's 5417 microprocessor-controlled, **benchtop centrifuge** features a brushless drive designed to permit short acceleration and deceleration times of 10 s and 12 s, respectively (*Reader Service No. 113*). The speed can be adjusted up to 14,000 r.p.m. (20,800 g). The 5417 accommodates a 24-place, fixed-angle rotor or a 12-place, swing-bucket rotor. Both rotors can be autoclaved and are designed for test tubes of up to 2 ml. As a safety feature, the aerosol-tight version of the fixed-angle rotor allows safe centrifugation of toxic, infectious or radioactive material. This model will operate in cold-rooms at +4 °C.

Millipore's Ultrafree-20, ready-to-use **centrifugal filters** are designed for the concentration, purification and desalting of biological samples (from 0.5 ml to 20 ml) (*Reader Service No. 114*). Depending on the starting concentration, 20 ml of sample can be concentrated down to 0.5 ml in 10–60 min, the company says. The filters are also recommended for a variety of other uses: the recovery of enzymes, virus, bacteria, yeast and peptides from biological samples; for concentrating serum or urine proteins for electrophoresis; and for the recovery of low-molecular-weight components from cell



**Ultrafree-20 centrifugal filters.**

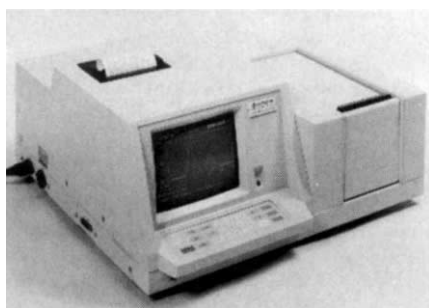
lysates, culture media, or fermentation broths. Each filter incorporates a PLGC ultrafiltration membrane (10,000 nominal molecular weight limit) made of low-binding regenerated cellulose. These devices can be used with standard swinging-bucket (horizontal) rotors for 50-ml centrifuge tubes.

Qiagen **genomic tips** are now available for small-to-large-scale preparation of genomic DNA. **NATURE** · VOL 371 · 20 OCTOBER 1994

nomic DNA from blood, cultured cells, tissue, yeast and Gram-negative bacteria (*Reader Service No. 115*). The procedure requires no phenol or chloroform. DNA purified in this way ranges in size from 20 to 150 kilobases, with most of the DNA lying in the 50–100-kb range. It is said to be free of contaminants such as RNA, protein and metabolites, with  $A_{260/280}$  ratios of 1.7–1.9.

### ■ Lab equipment

Hitachi's U-2000 **ultraviolet/visible spectrophotometer** has received a major software upgrade (*Reader Service No. 116*). Enhancements to the processing of calibration data and the generation of calibration curves have been incorporated to improve the quality of sample concentration measurements. Additionally, there are new application programs for the measurement of proteins and nucleic acid, and for automatic control of a sample sipper and six-cell positioner. The U-2000 now has the ability to generate a working curve using as many as 20 standards measured up to five times each to give increased accuracy and precision. These new standard measurements can



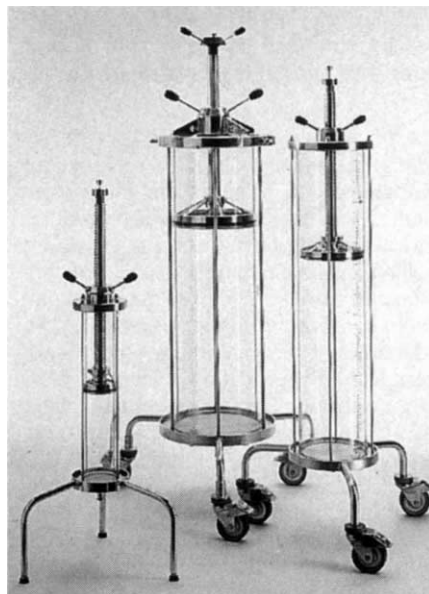
Upgraded U-2000 spectrophotometer.

now be fitted to linear or quadratic working curves; advanced statistical verification of the calibration curve gives the user greater confidence in the measurement of unknown samples by automatically indicating the quality of fit of the calculated working curve to the measured data. Data display facilities have been enhanced. Up to 13 spectra may be stored in nonvolatile memory and repeat scans can be overlaid for direct comparison.

Dionex has added the PD40 **diode-array detector** to its DX500 HPLC product line (*Reader Service No. 117*). The detector features a 512-element diode array with a 195–600 nm wavelength range, a temperature-controlled optics bench and high-intensity deuterium lamp. All functions, from control of the instrument through to interpretation of the data and reporting, are controlled with the accompanying Windows-based software. Data can be viewed in real-time as ratio plots, rotatable three-dimensional contour plots, multiple chromatograms, or a combination of these views. Data interpretation features include multi-wavelength analyses, peak-purity determinations and spectral library searches.

NATURE • VOL 371 • 20 OCTOBER 1994

BPG **columns** from Pharmacia Biotech are now available with pressure ratings from 4 to 8 bar, allowing the use of a wider range of chromatographic media — particularly



Columns with 4 to 8 bar pressure ratings.

the smaller particle size media (34  $\mu\text{m}$ ) for higher resolution (*Reader Service No. 118*). The columns are manufactured from calibrated precision glass and electropolished steel to hinder microbial attachment and growth. The single-screw adaptor helps in packing and a range of sizes is available.

### ■ Reagents

GlycoShift from Oxford GlycoSystems uses a **recombinant peptide-N-glycosidase F** (PNGase F) selectively to release N-glycans from the sample glycoprotein (*Reader Service No. 119*). A protocol supplied with the kit can also be used to determine the approximate number of glycan chains present. The rPNGase F used in the kit (which is also sold separately) is cloned from *Flavobacterium meningosepticum* and over-expressed in *Escherichia coli*.

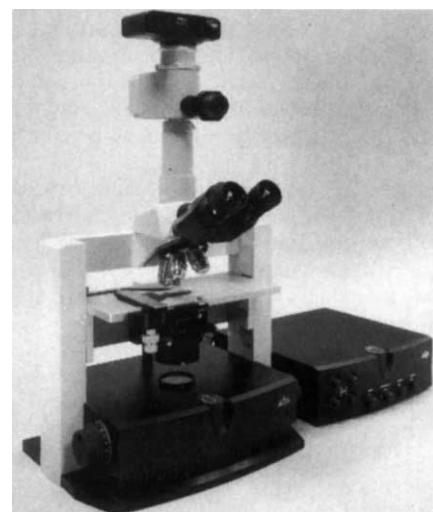
A **catalogue** of products and services is now available from Wallac and Berthold, EG & G companies that specialize in biospecific assays involving labels (*Reader Service No. 120*). Entitled 'Measurements With Light', the catalogue contains descriptions of both companies' instruments, software and reagents, with reference tables.

### ■ Microscopy matters

The **scientific workstation** from Melles Griot can be used for supporting patch-clamp apparatus, cell injection and micro-manipulation equipment, interferometers, microscopes, noncontact surface measurement devices and other motion-sensitive instruments (*Reader Service No. 121*). The basic workstation (which has a honeycomb-core construction) consists of a welded steel frame, with a self-levelling pneumatic damp-

ing system designed to provide both vertical and horizontal isolation for the 60-mm thick instrument platform.

Edge Scientific Instruments has launched a new **high-definition, real-time three-dimensional microscope**, the Edge R400, which should be of interest to researchers in both the biological and material sciences (*Reader Service No. 122*). The R400 features the company's Multiple Oblique illumination, consisting of four, independently adjustable illuminating beams arranged as left/right pairs. The beams pass through specimens at an oblique angle, generating separate left- and right-eye views. The resulting 3-D images can be observed directly



Edge R400 real-time 3D microscope.

through the eyepieces, photographed as stereo pairs or single-picture 3-D anaglyphs (colour-coded), and acquired via video for computer image processing and quantitative analysis. The Multiple Oblique illumination is designed to improve resolution, sharpness, contrast and depth of field — whether a specimen is viewed in three dimensions or two. □

These notes are compiled by Diane Gershon from information provided by the manufacturers. For more details, fill in the reader service card bound inside the journal.

### ADVERTISEMENT

## OLIGONUCLEOTIDES

Synthesised, purified and shipped in 3 days

**£2.50 per base**

(All inclusive)

**0.2  $\mu\text{M}$  scale (approx. 500  $\mu\text{g}$ )**

- Antisense (completely or partially sulphurised)
- Extensive range of labelling and modifications
- Full design and analysis facilities
- Gene cassettes, bulk order and contract discounts

**Molecular Medicine Unit**

King's College School of Medicine & Dentistry  
Tel: 071 346 3126 Fax: 071-733 3877

Reader Service No.10



# Future of the Euroscientist

David Collingham

**Labour mobility within the pharmaceutical and biotech sectors in Europe is becoming a crucial issue to both employers and employees.**

EARLIER in this series (see *Nature* 368, 778; 1994), I shared my experiences on career development among scientists, gained from over a decade of recruiting experience in the biosciences.

Within the many industrial R&D establishments in the UK and continental Europe, it is not uncommon to find employees of many nationalities working together at the bench and in management. This leads to the question of how transferable are one's scientific and management skills? Indeed, how transferable is a given individual?

A recent survey by accountants Ernst and Young and the SAGB (Senior Advisory Group on Biotechnology)<sup>1</sup>, relating to Europe's biotechnology industry, shows that the single most important factor influencing investment decisions of biotech companies is the availability of skilled staff in the locations chosen for their operation. They reckon the number of people working in European organisations depending on the application of biotechnology is at least 184,000, and that the annual rate of increase for small and medium enterprises is 6.5%. Even this pales, however, in the face of a predicted annual rate of 23% in the USA. Clearly, if these predictions are realised, the mobility of the European scientific labour market is about to be tested.

I recently questioned a number of human resource professionals in the global or European headquarters of world-ranked pharmaceutical or bioscience companies. Their answers varied, but the common theme was of a significant difference between theory and practice. In theory, employees within the EC can rotate jobs or transfer with ease between countries, aided by their laboratories or companies. Nothing could be simpler. In practice all sorts of problems raise their heads.

Location issues can be looked at in two distinct ways: through the eyes of the employer and the employee. For employers — pharmaceutical and biotech R&D organisations, for example — the choice of a new location is clearly complex and depends upon a myriad of intertwining factors. In Europe, many R&D centres are set up close to academic research centres; in the biosciences, the favourite locations seem to be Oxford and Cambridge in the UK, Heidelberg and Munich in Germany, and Leiden in The Netherlands, to name just a few. Foreign employ-

ing organisations, such as US or Japanese companies, will look at the many economic, scientific and social variables before deciding on a location. There are now in Europe a number of specialist consultancies dedicated to providing information to companies in these situations.

In order to attract senior R&D managers to a given location, there needs to be not only a favourable scientific climate but also a user-friendly social and economic environment. The tabular material in ref. 2 gives some broad idea of the varying combinations of compensation benefits around Europe, which may influence the employer's decision on the optimal location.

Turning to employees, what are the major factors affecting their job mobility? Much depends on the support given by the employer, of course, for this can have a crucial bearing on the ultimate success of the job placement. Many of the larger organisations have formal schemes in place — trans-Atlantic job swaps, job rotation, overseas secondment, project or task force placement, and so on. Thus a Briton arriving in, say, Glaxo's Italian R&D organisation in Verona is likely to be given a helping hand with the process of easing himself or herself into the new culture. Nevertheless the cultural gaps can be immense and the person appointed must learn not to expect favoured treatment and to become familiar with the local habits! When in Rome....

## Mobility problems

In our role as recruiters involved in helping to move individuals and their families across country borders, we see successes and some failures. Too often the failures arise from shortsightedness, narrowmindedness and a lack of appreciation of the implications of a family move. Conversely, an otherwise suitable candidate may not fully appreciate the upside of a career-enhancing transnational move, preferring to maintain the status quo rather than risk a major change.

A few years ago, we conducted a short survey of European biotechnology companies, the results of which showed a remarkably xenophobic attitude to labour mobility. Nowadays attitudes are changing and today's scientists and R&D managers are expected to be mobile, adaptable and flexible. I recently heard it remarked that the common language of science is broken English, a reflection not on educational

standards but on the ever-increasing panoply of nationalities found in the modern research laboratory.

The scenario of a Swiss-based pharmaceutical clinical research department moving *en bloc* to Strasbourg or south-east England is a reality and there may be more examples to follow. This means that the scientist who actively manages his/her career may need to show an increasing adaptability and flexibility to the possibilities of new cultural and geographical changes in job.

So for those readers contemplating a major move of this kind, there are a few hints and tips that will help career advancement and successful relocation:

■ Is the job right? Are your motives for moving the correct ones?

■ What are the career prospects? Is there a danger of being 'stuck out on a limb'? How easy will it be to reestablish yourself in your home country?

■ How well will you and your family adapt to your host country? Key factors include ability or need for spouse to work (for example, a Green Card is necessary in the USA), social integration within the host country, ability to learn the host language, and the extent to which your education ambitions for children can be met by the local system.

■ Local economic factors. How high are income taxation and other taxes? How does the cost of living compare with the home country? What is the availability of tax incentives (for example, in The Netherlands). What are pension arrangements like?

■ Frustration factors that need to be taken into account include local bureaucracy, and difficulty in obtaining housing and essential items such as telephones.

■ What are communications like from host country to home country? What is the distance and ease of travel to relatives and friends? □

*David Collingham is Chairman and Managing Partner of Ruston Poole International, founder member of Global Partners, a worldwide executive search network specialising in healthcare and biosciences. Ruston Poole International is based at 11/12 Buckingham Gate, London SW1E 6LB.*

1. *Biotechnology's Economic Impact in Europe* (Ernst & Young/SAGB, 1994).

2. *Pay and Benefits for Staff in Europe* (Collingham, D. P. & Grindley, J. *Genetic Engineering News*, 14; April 1994).